09/674546

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/31, C07K 14/22, 16/12, C12Q 1/68, A61K 39/095, G01N 33/50	A2	(11) International Publication Number: WO 99/57280 (43) International Publication Date: 11 November 1999 (11.11.99)

(21) International Application Number: PCT/US99/09346

(22) International Filing Date: 30 April 1999 (30.04.99)

(30) Priority Data:

60/083,758	1 May 1998 (01.05.98)	US
60/094,869	31 July 1998 (31.07.98)	US
60/098,994	2 September 1998 (02.09.98)	US
60/099,062	2 September 1998 (02.09.98)	US
60/103,749	9 October 1998 (09.10.98)	US
60/103,794	9 October 1998 (09.10.98)	US
60/103,796	9 October 1998 (09.10.98)	US
60/121,528	25 February 1999 (25.02.99)	US

(71) Applicants *(for all designated States except US)*: CHIRON CORPORATION [US/US]; 4560 Horton Street, Emeryville, CA 94608 (US). THE INSTITUTE FOR GENOMIC RESEARCH [US/US]; 9212 Medical Center Drive, Rockville, MD 20850 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: FRASER, Claire [US/US]; Rockville, MD (US). GALEOTTI, Cesira [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). GRANDI, Guido [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). HICKEY, Erin [US/US]; Gaithersburg, MD (US). MASIGNANI, Vega [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). MORA, Marirosa [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). PETERSEN, Jeremy [US/US]; Arlington, VA (US). PIZZA, Mariagratia [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). RAPPUOLI, Rino [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). RATTI, Giulio [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). SCALATO, Enzo [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). SCARSELLI, Maria [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). TETTELIN, Herve [US/US]; Gaithersburg, MD (US). VENTER, J., Craig [US/US]; Rockville, MD (US).

(74) Agent: HARBIN, Alisa, A.; Chiron Corporation, Intellectual Property – R440, P.O. Box 8097, Emeryville, CA 94662-8097 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: NEISSERIA MENINGITIDIS ANTIGENS AND COMPOSITIONS

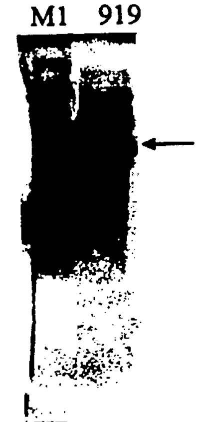
(57) Abstract

The invention provides proteins from Neisseria meningitidis, including the amino acid sequences and the corresponding nucleotide sequences. The proteins are predicted to be useful antigens for vaccines and/or diagnostics.

919 (46 kDa)

A) PURIFICATION

M1 919



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

NEISSERIA MENINGITIDIS ANTIGENS AND COMPOSITIONS

This application is a continuation-in-part of the following U.S. Provisional Patent applications, from each of which priority is claimed, and each of which is incorporated by reference in its entirety: 60/083,758 (filed May 1, 1998); 60/094,869 (filed July 31, 1998); 60/098,994 (filed September 2, 1998); 60/099,062 (filed September 2, 1998); 60/103,749 (filed October 9, 1998); 60/103,794 (filed October 9, 1998); 60/103,796 (filed October 9, 1998); and 60/121,528 (filed February 25. 1999).

This invention relates to antigens from the bacterial species: *Neisseria meningitidis* and *Neisseria gonorrhoeae.*

BACKGROUND

Neisseria meningitidis is a non-motile, gram negative diplococcus human pathogen. It colonizes the pharynx, causing meningitis and, occasionally, septicaemia in the absence of meningitis. It is closely related to *N. gonorrhoea*. although one feature that clearly differentiates meningococcus from gonococcus is the presence of a polysaccharide capsule that is present in all pathogenic meningococci.

N. meningitidis causes both endemic and epidemic disease. In the United States the attack rate is 0.6-1 per 100,000 persons per year, and it can be much greater during outbreaks. (see Lieberman *et al.* (1996) Safety and Immunogenicity of a Serogroups A/C *Neisseria meningitidis* Oligosaccharide-Protein Conjugate Vaccine in Young Children. *JAMA* 275(19):1499-1503; Schuchat *et al* (1997) Bacterial Meningitis in the United States in 1995. *N Engl J Med* 337(14):970-976). In developing countries, endemic disease rates are much higher and during epidemics incidence rates can reach 500 cases per 100,000 persons per year. Mortality is extremely high, at 10-20% in the United States, and much higher in developing countries. Following the introduction of the conjugate vaccine against *Haemophilus influenzae*, *N. meningitidis* is the major cause of bacterial meningitis at all ages in the United States (Schuchat *et al* (1997) *supra*).

Based on the organism's capsular polysaccharide, 12 serogroups of *N. meningitidis* have been identified. Group A is the pathogen most often implicated in epidemic disease in sub-Saharan Africa. Serogroups B and C are responsible for the vast majority of cases in the

United States and in most developed countries. Serogroups W135 and Y are responsible for the rest of the cases in the United States and developed countries. The meningococcal vaccine currently in use is a tetravalent polysaccharide vaccine composed of serogroups A, C, Y and W135. Although efficacious in adolescents and adults, it induces a poor immune response and short duration of protection, and cannot be used in infants [*eg.* Morbidity and Mortality weekly report, Vol.46, No. RR-5 (1997)]. This is because polysaccharides are T-cell independent antigens that induce a weak immune response that cannot be boosted by repeated immunization. Following the success of the vaccination against *H.influenzae*, conjugate vaccines against serogroups A and C have been developed and are at the final stage of clinical testing (Zollinger WD "New and Improved Vaccines Against Meningococcal Disease". In: *New Generation Vaccines*, *supra*, pp. 469-488; Lieberman *et al* (1996) *supra*; Costantino *et al* (1992) Development and phase I clinical testing of a conjugate vaccine against meningococcus A and C. *Vaccine* 10:691-698).

Meningococcus B (menB) remains a problem, however. This serotype currently is responsible for approximately 50% of total meningitis in the United States, Europe, and South America. The polysaccharide approach cannot be used because the menB capsular polysaccharide is a polymer of α(2-8)-linked *N*-acetyl neuraminic acid that is also present in mammalian tissue. This results in tolerance to the antigen; indeed, if an immune response were elicited, it would be anti-self, and therefore undesirable. In order to avoid induction of autoimmunity and to induce a protective immune response, the capsular polysaccharide has, for instance, been chemically modified substituting the *N*-acetyl groups with *N*-propionyl groups, leaving the specific antigenicity unaltered (Romero & Outschoorn (1994) Current status of Meningococcal group B vaccine candidates: capsular or non-capsular? *Clin Microbiol Rev* 7(4):559-575).

Alternative approaches to menB vaccines have used complex mixtures of outer membrane proteins (OMPs), containing either the OMPs alone, or OMPs enriched in porins, or deleted of the class 4 OMPs that are believed to induce antibodies that block bactericidal activity. This approach produces vaccines that are not well characterized. They are able to protect against the homologous strain, but are not effective at large where there are many antigenic variants of the outer membrane proteins. To overcome the antigenic variability, multivalent vaccines containing up to nine different porins have been constructed (*eg.* Poolman JT (1992) Development of a meningococcal vaccine. *Infect. Agents Dis.* 4:13-28).

Additional proteins to be used in outer membrane vaccines have been the opa and opc proteins, but none of these approaches have been able to overcome the antigenic variability (*eg.* Ala'Aldeen & Borriello (1996) The meningococcal transferrin-binding proteins 1 and 2 are both surface exposed and generate bactericidal antibodies capable of killing homologous and heterologous strains. *Vaccine* 14(1):49-53).

A certain amount of sequence data is available for meningococcal and gonoccocal genes and proteins (*e.g.* EP-A-0467714, WO96/29412), but this is by no means complete. Other men B proteins may include those listed in WO 97/28273, WO 96/29412, WO 95/03413, US 5,439,808, and US 5,879,686.

The provision of further sequences could provide an opportunity to identify secreted or surface-exposed proteins that are presumed targets for the immune system and which are not antigenically variable. For instance, some of the identified proteins could be components of efficacious vaccines against meningococcus B, some could be components of vaccines against all meningococcal serotypes, and others could be components of vaccines against all pathogenic *Neisseriae* including *Neisseria meningitidis* or *Neisseria gonorrhoeae*. Those sequences specific to *N. meningitidis* or *N. gonorrhoeae* that are more highly conserved are further preferred sequences.

It is thus an object of the invention is to provide Neisserial DNA sequences which encode proteins that are antigenic or immunogenic.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 illustrates the products of protein expression and purification of the predicted ORF 919 as cloned and expressed in *E. coli.*

Fig. 2 illustrates the products of protein expression and purification of the predicted ORF 279 as cloned and expressed in *E. coli.*

Fig. 3 illustrates the products of protein expression and purification of the predicted ORF 576-1 as cloned and expressed in *E. coli.*

Fig. 4 illustrates the products of protein expression and purification of the predicted ORF 519-1 as cloned and expressed in *E. coli.*

Fig. 5 illustrates the products of protein expression and purification of the predicted ORF 121-1 as cloned and expressed in *E. coli.*

Fig. 6 illustrates the products of protein expression and purification of the predicted ORF 128-1 as cloned and expressed in *E. coli*.

Fig. 7 illustrates the products of protein expression and purification of the predicted ORF 206 as cloned and expressed in *E. coli*.

Fig. 8 illustrates the products of protein expression and purification of the predicted ORF 287 as cloned and expressed in *E. coli*.

Fig. 9 illustrates the products of protein expression and purification of the predicted ORF 406 as cloned and expressed in *E. coli*.

Fig. 10 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 919 as cloned and expressed in *E. coli*.

Fig. 11 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 279 as cloned and expressed in *E. coli*.

Fig. 12 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 576-1 as cloned and expressed in *E. coli*.

Fig. 13 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 519-1 as cloned and expressed in *E. coli*.

Fig. 14 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 121-1 as cloned and expressed in *E. coli*.

Fig. 15 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 128-1 as cloned and expressed in *E. coli*.

Fig. 16 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 206 as cloned and expressed in *E. coli*.

Fig. 17 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 287 as cloned and expressed in *E. coli*.

Fig. 18 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 406 as cloned and expressed in *E. coli*.

Fig. 19 shows an alignment comparison of amino acid sequences for ORF 225 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

Fig. 20 shows an alignment comparison of amino acid sequences for ORF 235 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

Fig. 21 shows an alignment comparison of amino acid sequences for ORF 287 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

Fig. 22 shows an alignment comparison of amino acid sequences for ORF 519 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

Fig. 23 shows an alignment comparison of amino acid sequences for ORF 919 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

THE INVENTION

The invention provides proteins comprising the *N. meningitidis* amino acid sequences and *N. gonorrhoeae* amino acid sequences disclosed in the examples.

It also provides proteins comprising sequences homologous (i.e., those having sequence identity) to the *N. meningitidis* amino acid sequences disclosed in the examples. Depending on the particular sequence, the degree of homology (sequence identity) is preferably greater than 50% (*eg*. 60%, 70%, 80%, 90%, 95%, 99% or more). These proteins include mutants and allelic variants of the sequences disclosed in the examples. Typically, 50% identity or more between two proteins is considered to be an indication of functional equivalence. Identity between proteins is preferably determined by the Smith-Waterman

homology search algorithm as implemented in MPSRCH program (Oxford Molecular) using an affine gap search with parameters:gap penalty 12, gap extension penalty 1.

The invention further provides proteins comprising fragments of the *N. meningitidis* amino acid sequences and *N. gonorrhoeae* amino acid sequences disclosed in the examples. The fragments should comprise at least *n* consecutive amino acids from the sequences and, depending on the particular sequence, *n* is 7 or more (*eg.* 8, 10, 12, 14, 16, 18, 20 or more). Preferably the fragments comprise an epitope from the sequence.

The proteins of the invention can, of course, be prepared by various means (*eg.* recombinant expression, purification from cell culture, chemical synthesis *etc.*) and in various forms (*eg.* native, fusions *etc.*). They are preferably prepared in substantially pure or isolated form (*ie.* substantially free from other *N. meningitidis* or *N. gonorrhoeae* host cell proteins)

According to a further aspect, the invention provides antibodies which bind to these proteins. These may be polyclonal or monoclonal and may be produced by any suitable means.

According to a further aspect, the invention provides nucleic acid comprising the *N. meningitidis* nucleotide sequences and *N. gonorrhoeae* nucleotide sequences disclosed in the examples.

According to a further aspect, the invention comprises nucleic acids having sequence identity of greater than 50% (e.g., 60%, 70%, 80%, 90%, 95%, 99% or more) to the nucleic acid sequences herein. Sequence identity is determined as above-discussed.

According to a further aspect, the invention comprises nucleic acid that hybridizes to the sequences provided herein. Conditions for hybridization are set forth herein.

Nucleic acid comprising fragments of these sequences are also provided. These should comprise at least *n* consecutive nucleotides from the *N. meningitidis* sequences or *N. gonorrhoeae* sequences and depending on the particular sequence, *n* is 10 or more (*eg* 12, 14, 15, 18, 20, 25, 30, 35, 40 or more).

According to a further aspect, the invention provides nucleic acid encoding the proteins and protein fragments of the invention.

It should also be appreciated that the invention provides nucleic acid comprising sequences complementary to those described above (*eg.* for antisense or probing purposes).

Nucleic acid according to the invention can, of course, be prepared in many ways (*eg.* by chemical synthesis, in part or in whole, from genomic or cDNA libraries, from the

organism itself *etc.*) and can take various forms (*eg.* single stranded, double stranded, vectors, probes *etc.*).

In addition, the term "nucleic acid" includes DNA and RNA, and also their analogues, such as those containing modified backbones, and also protein nucleic acids (PNA) *etc.*

According to a further aspect, the invention provides vectors comprising nucleotide sequences of the invention (*eg.* expression vectors) and host cells transformed with such vectors.

According to a further aspect, the invention provides compositions comprising protein, antibody, and/or nucleic acid according to the invention. These compositions may be suitable as vaccines, for instance, or as diagnostic reagents or as immunogenic compositions.

The invention also provides nucleic acid, protein, or antibody according to the invention for use as medicaments (*eg.* as vaccines) or as diagnostic reagents. It also provides the use of nucleic acid, protein, or antibody according to the invention in the manufacture of (I) a medicament for treating or preventing infection due to Neisserial bacteria (ii) a diagnostic reagent for detecting the presence of Neisserial bacteria or of antibodies raised against Neisserial bacteria or (iii) for raising antibodies. Said Neisserial bacteria may be any species or strain (such as *N. gonorrhoeae*) but are preferably *N. meningitidis*, especially strain B or strain C.

The invention also provides a method of treating a patient, comprising administering to the patient a therapeutically effective amount of nucleic acid, protein, and/or antibody according to the invention.

According to further aspects, the invention provides various processes.

A process for producing proteins of the invention is provided, comprising the step of culturing a host cell according to the invention under conditions which induce protein expression.

A process for detecting polynucleotides of the invention is provided, comprising the steps of: (a) contacting a nucleic probe according to the invention with a biological sample under hybridizing conditions to form duplexes; and (b) detecting said duplexes.

A process for detecting proteins of the invention is provided, comprising the steps of: (a) contacting an antibody according to the invention with a biological sample under conditions suitable for the formation of an antibody-antigen complexes; and (b) detecting said complexes.

Having now generally described the inventiion, the same will be more readily understood through reference to the following examples which are provided by way of illustration, and are not intended to be limiting of the present invention, unless specified.

Methodology - Summary of standard procedures and techniques.

General

This invention provides *Neisseria meningitidis* menB nucleotide sequences, amino acid sequences encoded therein. With these disclosed sequences, nucleic acid probe assays and expression cassettes and vectors can be produced. The expression vectors can be transformed into host cells to produce proteins. The purified or isolated polypeptides (which may also be chemically synthesized) can be used to produce antibodies to detect menB proteins. Also, the host cells or extracts can be utilized for biological assays to isolate agonists or antagonists. In addition, with these sequences one can search to identify open reading frames and identify amino acid sequences. The proteins may also be used in immunogenic compositions, antigenic compositions and as vaccine components.

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature e.g., Sambrook *Molecular Cloning; A Laboratory Manual, Second Edition* (1989); *DNA Cloning, Volumes I and ii* (D.N Glover ed. 1985); *Oligonucleotide Synthesis* (M.J. Gait ed, 1984); *Nucleic Acid Hybridization* (B.D. Hames & S.J. Higgins eds. 1984); *Transcription and Translation* (B.D. Hames & S.J. Higgins eds. 1984); *Animal Cell Culture* (R.I. Freshney ed. 1986); *Immobilized Cells and Enzymes* (IRL Press, 1986); B. Perbal, *A Practical Guide to Molecular Cloning* (1984); the *Methods in Enzymology* series (Academic Press, Inc.), especially volumes 154 & 155; *Gene Transfer Vectors for Mammalian Cells* (J.H. Miller and M.P. Calos eds. 1987, Cold Spring Harbor Laboratory); Mayer and Walker, eds. (1987), *Immunochemical Methods in Cell and Molecular Biology* (Academic Press, London); Scopes, (1987) *Protein Purification: Principles and Practice*, Second Edition (Springer-Verlag, N.Y.), and *Handbook of Experimental Immunology, Volumes I-IV* (D.M. Weir and C.C. Blackwell eds 1986).

Standard abbreviations for nucleotides and amino acids are used in this specification.

All publications, patents, and patent applications cited herein are incorporated in full by reference.

Expression systems

The *Neisseria* menB nucleotide sequences can be expressed in a variety of different expression systems; for example those used with mammalian cells, plant cells, baculoviruses, bacteria, and yeast.

i. Mammalian Systems

Mammalian expression systems are known in the art. A mammalian promoter is any DNA sequence capable of binding mammalian RNA polymerase and initiating the downstream (3') transcription of a coding sequence (e.g., structural gene) into mRNA. A promoter will have a transcription initiating region, which is usually placed proximal to the 5' end of the coding sequence, and a TATA box, usually located 25-30 base pairs (bp) upstream of the transcription initiation site. The TATA box is thought to direct RNA polymerase II to begin RNA synthesis at the correct site. A mammalian promoter will also contain an upstream promoter element, usually located within 100 to 200 bp upstream of the TATA box. An upstream promoter element determines the rate at which transcription is initiated and can act in either orientation (Sambrook et al. (1989) "Expression of Cloned Genes in Mammalian Cells." In *Molecular Cloning: A Laboratory Manual, 2nd ed.).*

Mammalian viral genes are often highly expressed and have a broad host range; therefore sequences encoding mammalian viral genes provide particularly useful promoter sequences. Examples include the SV40 early promoter, mouse mammary tumor virus LTR promoter, adenovirus major late promoter (Ad MLP), and herpes simplex virus promoter. In addition, sequences derived from non-viral genes, such as the murine metallothionein gene, also provide useful promoter sequences. Expression may be either constitutive or regulated (inducible). Depending on the promoter selected, many promotes may be inducible using known substrates, such as the use of the mouse mammary tumor virus (MMTV) promoter with the glucocorticoid responsive element (GRE) that is induced by glucocorticoid in hormone-responsive transformed cells (see for example, U.S. Patent 5,783,681).

The presence of an enhancer element (enhancer), combined with the promoter elements described above, will usually increase expression levels. An enhancer is a

regulatory DNA sequence that can stimulate transcription up to 1000-fold when linked to homologous or heterologous promoters, with synthesis beginning at the normal RNA start site. Enhancers are also active when they are placed upstream or downstream from the transcription initiation site, in either normal or flipped orientation, or at a distance of more than 1000 nucleotides from the promoter (Maniatis et al. (1987) *Science 236:*1237; Alberts et al. (1989) *Molecular Biology of the Cell*, 2nd ed.). Enhancer elements derived from viruses may be particularly useful, because they usually have a broader host range. Examples include the SV40 early gene enhancer (Dijkema et al (1985) *EMBO J. 4*:761) and the enhancer/promoters derived from the long terminal repeat (LTR) of the Rous Sarcoma Virus (Gorman et al. (1982b) *Proc. Natl. Acad. Sci. 79*:6777) and from human cytomegalovirus (Boshart et al. (1985) *Cell 41*:521). Additionally, some enhancers are regulatable and become active only in the presence of an inducer, such as a hormone or metal ion (Sassone-Corsi and Borelli (1986) *Trends Genet. 2*:215; Maniatis et al. (1987) Science 236:1237).

A DNA molecule may be expressed intracellularly in mammalian cells. A promoter sequence may be directly linked with the DNA molecule, in which case the first amino acid at the N-terminus of the recombinant protein will always be a methionine, which is encoded by the ATG start codon. If desired, the N-terminus may be cleaved from the protein by *in vitro* incubation with cyanogen bromide.

Alternatively, foreign proteins can also be secreted from the cell into the growth media by creating chimeric DNA molecules that encode a fusion protein comprised of a leader sequence fragment that provides for secretion of the foreign protein in mammalian cells. Preferably, there are processing sites encoded between the leader fragment and the foreign gene that can be cleaved either *in vivo* or *in vitro*. The leader sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell. The adenovirus tripartite leader is an example of a leader sequence that provides for secretion of a foreign protein in mammalian cells.

Usually, transcription termination and polyadenylation sequences recognized by mammalian cells are regulatory regions located 3' to the translation stop codon and thus, together with the promoter elements, flank the coding sequence. The 3' terminus of the mature mRNA is formed by site-specific post-transcriptional cleavage and polyadenylation (Birnstiel et al. (1985) *Cell 41*:349; Proudfoot and Whitelaw (1988) "Termination and 3' end processing of eukaryotic RNA. In *Transcription and splicing* (ed. B.D. Hames and D.M.

Glover); Proudfoot (1989) *Trends Biochem. Sci. 14*:105). These sequences direct the transcription of an mRNA which can be translated into the polypeptide encoded by the DNA. Examples of transcription terminator/polyadenylation signals include those derived from SV40 (Sambrook et al (1989) "Expression of cloned genes in cultured mammalian cells." In *Molecular Cloning: A Laboratory Manual*).

Usually, the above described components, comprising a promoter, polyadenylation signal, and transcription termination sequence are put together into expression constructs. Enhancers, introns with functional splice donor and acceptor sites, and leader sequences may also be included in an expression construct, if desired. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as mammalian cells or bacteria. Mammalian replication systems include those derived from animal viruses, which require trans-acting factors to replicate. For example, plasmids containing the replication systems of papovaviruses, such as SV40 (Gluzman (1981) *Cell 23*:175) or polyomavirus, replicate to extremely high copy number in the presence of the appropriate viral T antigen. Additional examples of mammalian replicons include those derived from bovine papillomavirus and Epstein-Barr virus. Additionally, the replicon may have two replication systems, thus allowing it to be maintained, for example, in mammalian cells for expression and in a prokaryotic host for cloning and amplification. Examples of such mammalian-bacteria shuttle vectors include pMT2 (Kaufman et al. (1989) *Mol. Cell. Biol. 9*:946) and pHEBO (Shimizu et al. (1986) *Mol. Cell. Biol. 6*:1074).

The transformation procedure used depends upon the host to be transformed. Methods for introduction of heterologous polynucleotides into mammalian cells are known in the art and include dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei.

Mammalian cell lines available as hosts for expression are known in the art and include many immortalized cell lines available from the American Type Culture Collection (ATCC), including but not limited to, Chinese hamster ovary (CHO) cells, HeLa cells, baby hamster kidney (BHK) cells, monkey kidney cells (COS), human hepatocellular carcinoma cells (e.g., Hep G2), and a number of other cell lines.

ii. Plant Cellular Expression Systems

There are many plant cell culture and whole plant genetic expression systems known in the art. Exemplary plant cellular genetic expression systems include those described in patents, such as: U.S. 5,693,506; US 5,659,122; and US 5,608,143. Additional examples of genetic expression in plant cell culture has been described by Zenk, *Phytochemistry* 30:3861-3863 (1991). Descriptions of plant protein signal peptides may be found in addition to the references described above in Vaulcombe et al., *Mol. Gen. Genet.* 209:33-40 (1987); Chandler et al., *Plant Molecular Biology* 3:407-418 (1984); Rogers, *J. Biol. Chem.* 260:3731-3738 (1985); Rothstein et al., *Gene* 55:353-356 (1987); Whittier et al., Nucleic Acids Research 15:2515-2535 (1987); Wirsel et al., *Molecular Microbiology* 3:3-14 (1989); Yu et al., *Gene* 122:247-253 (1992). A description of the regulation of plant gene expression by the phytohormone, gibberellic acid and secreted enzymes induced by gibberellic acid can be found in R.L. Jones and J. MacMillin, Gibberellins: in: *Advanced Plant Physiology*,. Malcolm B. Wilkins, ed., 1984 Pitman Publishing Limited, London, pp. 21-52. References that describe other metabolically-regulated genes: Sheen, *Plant Cell*, 2:1027-1038(1990); Maas et al., *EMBO J.* 9:3447-3452 (1990); Benkel and Hickey, *Proc. Natl. Acad. Sci.* 84:1337-1339 (1987)

Typically, using techniques known in the art, a desired polynucleotide sequence is inserted into an expression cassette comprising genetic regulatory elements designed for operation in plants. The expression cassette is inserted into a desired expression vector with companion sequences upstream and downstream from the expression cassette suitable for expression in a plant host. The companion sequences will be of plasmid or viral origin and provide necessary characteristics to the vector to permit the vectors to move DNA from an original cloning host, such as bacteria, to the desired plant host. The basic bacterial/plant vector construct will preferably provide a broad host range prokaryote replication origin; a prokaryote selectable marker; and, for Agrobacterium transformations, T DNA sequences for Agrobacterium-mediated transfer to plant chromosomes. Where the heterologous gene is not readily amenable to detection, the construct will preferably also have a selectable marker gene suitable for determining if a plant cell has been transformed. A general review of suitable markers, for example for the members of the grass family, is found in Wilmink and Dons, 1993, *Plant Mol. Biol. Reptr*, 11(2):165-185.

Sequences suitable for permitting integration of the heterologous sequence into the plant genome are also recommended. These might include transposon sequences and the like for homologous recombination as well as Ti sequences which permit random insertion of a heterologous expression cassette into a plant genome. Suitable prokaryote selectable markers include resistance toward antibiotics such as ampicillin or tetracycline. Other DNA sequences encoding additional functions may also be present in the vector, as is known in the art.

The nucleic acid molecules of the subject invention may be included into an expression cassette for expression of the protein(s) of interest. Usually, there will be only one expression cassette, although two or more are feasible. The recombinant expression cassette will contain in addition to the heterologous protein encoding sequence the following elements, a promoter region, plant 5' untranslated sequences, initiation codon depending upon whether or not the structural gene comes equipped with one, and a transcription and translation termination sequence. Unique restriction enzyme sites at the 5' and 3' ends of the cassette allow for easy insertion into a pre-existing vector.

A heterologous coding sequence may be for any protein relating to the present invention. The sequence encoding the protein of interest will encode a signal peptide which allows processing and translocation of the protein, as appropriate, and will usually lack any sequence which might result in the binding of the desired protein of the invention to a membrane. Since, for the most part, the transcriptional initiation region will be for a gene which is expressed and translocated during germination, by employing the signal peptide which provides for translocation, one may also provide for translocation of the protein of interest. In this way, the protein(s) of interest will be translocated from the cells in which they are expressed and may be efficiently harvested. Typically secretion in seeds are across the aleurone or scutellar epithelium layer into the endosperm of the seed. While it is not required that the protein be secreted from the cells in which the protein is produced, this facilitates the isolation and purification of the recombinant protein.

Since the ultimate expression of the desired gene product will be in a eucaryotic cell it is desirable to determine whether any portion of the cloned gene contains sequences which will be processed out as introns by the host's splicosome machinery. If so, site-directed mutagenesis of the "intron" region may be conducted to prevent losing a portion of the genetic message as a false intron code, Reed and Maniatis, *Cell* 41:95-105, 1985.

The vector can be microinjected directly into plant cells by use of micropipettes to mechanically transfer the recombinant DNA. Crossway, *Mol. Gen. Genet*, 202:179-185, 1985. The genetic material may also be transferred into the plant cell by using polyethylene glycol, Krens, et al., *Nature*, 296, 72-74, 1982. Another method of introduction of nucleic acid segments is high velocity ballistic penetration by small particles with the nucleic acid either within the matrix of small beads or particles, or on the surface, Klein, et al., *Nature*, 327, 70-73, 1987 and Knudsen and Muller, 1991, *Planta*, 185:330-336 teaching particle bombardment of barley endosperm to create transgenic barley. Yet another method of introduction would be fusion of protoplasts with other entities, either minicells, cells, lysosomes or other fusible lipid-surfaced bodies, Fraley, et al., *Proc. Natl. Acad. Sci. USA*, 79, 1859-1863, 1982.

The vector may also be introduced into the plant cells by electroporation. (Fromm et al., *Proc. Natl Acad. Sci. USA* 82:5824, 1985). In this technique, plant protoplasts are electroporated in the presence of plasmids containing the gene construct. Electrical impulses of high field strength reversibly permeabilize biomembranes allowing the introduction of the plasmids. Electroporated plant protoplasts reform the cell wall, divide, and form plant callus.

All plants from which protoplasts can be isolated and cultured to give whole regenerated plants can be transformed by the present invention so that whole plants are recovered which contain the transferred gene. It is known that practically all plants can be regenerated from cultured cells or tissues, including but not limited to all major species of sugarcane, sugar beet, cotton, fruit and other trees, legumes and vegetables. Some suitable plants include, for example, species from the genera *Fragaria*, *Lotus*, *Medicago*, *Onobrychis*, *Trifolium*, *Trigonella*, *Vigna*, *Citrus*, *Linum*, *Geranium*, *Manihot*, *Daucus*, *Arabidopsis*, *Brassica*, *Raphanus*, *Sinapis*, *Atropa*, *Capsicum*, *Datura*, *Hyoscyamus*, *Lycopersion*, *Nicotiana*, *Solanum*, *Petunia*, *Digitalis*, *Majorana*, *Cichorium*, *Helianthus*, *Lactuca*, *Bromus*, *Asparagus*, *Antirrhinum*, *Hererocallis*, *Nemesia*, *Pelargonium*, *Panicum*, *Pennisetum*, *Ranunculus*, *Senecio*, *Salpiglossis*, *Cucumis*, *Browaalia*, *Glycine*, *Lolium*, *Zea*, *Triticum*, *Sorghum*, and *Datura*.

Means for regeneration vary from species to species of plants, but generally a suspension of transformed protoplasts containing copies of the heterologous gene is first provided. Callus tissue is formed and shoots may be induced from callus and subsequently rooted. Alternatively, embryo formation can be induced from the protoplast suspension.

These embryos germinate as natural embryos to form plants. The culture media will generally contain various amino acids and hormones, such as auxin and cytokinins. It is also advantageous to add glutamic acid and proline to the medium, especially for such species as corn and alfalfa. Shoots and roots normally develop simultaneously. Efficient regeneration will depend on the medium, on the genotype, and on the history of the culture. If these three variables are controlled, then regeneration is fully reproducible and repeatable.

In some plant cell culture systems, the desired protein of the invention may be excreted or alternatively, the protein may be extracted from the whole plant. Where the desired protein of the invention is secreted into the medium, it may be collected. Alternatively, the embryos and embryoless-half seeds or other plant tissue may be mechanically disrupted to release any secreted protein between cells and tissues. The mixture may be suspended in a buffer solution to retrieve soluble proteins. Conventional protein isolation and purification methods will be then used to purify the recombinant protein. Parameters of time, temperature pH, oxygen, and volumes will be adjusted through routine methods to optimize expression and recovery of heterologous protein.

iii. Baculovirus Systems

The polynucleotide encoding the protein can also be inserted into a suitable insect expression vector, and is operably linked to the control elements within that vector. Vector construction employs techniques which are known in the art. Generally, the components of the expression system include a transfer vector, usually a bacterial plasmid, which contains both a fragment of the baculovirus genome, and a convenient restriction site for insertion of the heterologous gene or genes to be expressed; a wild type baculovirus with a sequence homologous to the baculovirus-specific fragment in the transfer vector (this allows for the homologous recombination of the heterologous gene in to the baculovirus genome); and appropriate insect host cells and growth media.

After inserting the DNA sequence encoding the protein into the transfer vector, the vector and the wild type viral genome are transfected into an insect host cell where the vector and viral genome are allowed to recombine. The packaged recombinant virus is expressed and recombinant plaques are identified and purified. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *inter alia*, Invitrogen, San Diego CA ("MaxBac" kit). These techniques are generally known to

those skilled in the art and fully described in Summers and Smith, *Texas Agricultural Experiment Station Bulletin No. 1555* (1987) (hereinafter "Summers and Smith").

Prior to inserting the DNA sequence encoding the protein into the baculovirus genome, the above described components, comprising a promoter, leader (if desired), coding sequence of interest, and transcription termination sequence, are usually assembled into an intermediate transplacement construct (transfer vector). This construct may contain a single gene and operably linked regulatory elements; multiple genes, each with its owned set of operably linked regulatory elements; or multiple genes, regulated by the same set of regulatory elements. Intermediate transplacement constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as a bacterium. The replicon will have a replication system, thus allowing it to be maintained in a suitable host for cloning and amplification.

Currently, the most commonly used transfer vector for introducing foreign genes into AcNPV is pAc373. Many other vectors, known to those of skill in the art, have also been designed. These include, for example, pVL985 (which alters the polyhedrin start codon from ATG to ATT, and which introduces a BamHI cloning site 32 basepairs downstream from the ATT; see Luckow and Summers, *Virology* (1989) *17*:31.

The plasmid usually also contains the polyhedrin polyadenylation signal (Miller et al. (1988) *Ann. Rev. Microbiol.*, *42*:177) and a prokaryotic ampicillin-resistance (*amp*) gene and origin of replication for selection and propagation in *E. coli*.

Baculovirus transfer vectors usually contain a baculovirus promoter. A baculovirus promoter is any DNA sequence capable of binding a baculovirus RNA polymerase and initiating the downstream (5' to 3') transcription of a coding sequence (e.g., structural gene) into mRNA. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site and a transcription initiation site. A baculovirus transfer vector may also have a second domain called an enhancer, which, if present, is usually distal to the structural gene. Expression may be either regulated or constitutive.

Structural genes, abundantly transcribed at late times in a viral infection cycle, provide particularly useful promoter sequences. Examples include sequences derived from the gene encoding the viral polyhedron protein, Friesen et al., (1986) "The Regulation of Baculovirus Gene Expression," in: *The Molecular Biology of Baculoviruses* (ed. Walter Doerfler); EPO

Publ. Nos. 127 839 and 155 476; and the gene encoding the p10 protein, Vlak et al., (1988), *J. Gen. Virol. 69*:765.

DNA encoding suitable signal sequences can be derived from genes for secreted insect or baculovirus proteins, such as the baculovirus polyhedrin gene (Carbonell et al. (1988) *Gene*, *73*:409). Alternatively, since the signals for mammalian cell posttranslational modifications (such as signal peptide cleavage, proteolytic cleavage, and phosphorylation) appear to be recognized by insect cells, and the signals required for secretion and nuclear accumulation also appear to be conserved between the invertebrate cells and vertebrate cells, leaders of non-insect origin, such as those derived from genes encoding human (alpha) α-interferon, Maeda et al., (1985), *Nature 315*:592; human gastrin-releasing peptide, Lebacq-Verheyden et al., (1988), *Molec. Cell. Biol. 8*:3129; human IL-2, Smith et al., (1985) *Proc. Nat'l Acad. Sci. USA*, *82*:8404; mouse IL-3, (Miyajima et al., (1987) *Gene 58*:273; and human glucocerebrosidase, Martin et al. (1988) *DNA*, *7*:99, can also be used to provide for secretion in insects.

A recombinant polypeptide or polyprotein may be expressed intracellularly or, if it is expressed with the proper regulatory sequences, it can be secreted. Good intracellular expression of nonfused foreign proteins usually requires heterologous genes that ideally have a short leader sequence containing suitable translation initiation signals preceding an ATG start signal. If desired, methionine at the N-terminus may be cleaved from the mature protein by *in vitro* incubation with cyanogen bromide.

Alternatively, recombinant polyproteins or proteins which are not naturally secreted can be secreted from the insect cell by creating chimeric DNA molecules that encode a fusion protein comprised of a leader sequence fragment that provides for secretion of the foreign protein in insects. The leader sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the translocation of the protein into the endoplasmic reticulum.

After insertion of the DNA sequence and/or the gene encoding the expression product precursor of the protein, an insect cell host is co-transformed with the heterologous DNA of the transfer vector and the genomic DNA of wild type baculovirus -- usually by co-transfection. The promoter and transcription termination sequence of the construct will usually comprise a 2-5kb section of the baculovirus genome. Methods for introducing heterologous DNA into the desired site in the baculovirus virus are known in the art. (See

Summers and Smith *supra*; Ju et al. (1987); Smith et al., *Mol. Cell. Biol.* (1983) *3*:2156; and Luckow and Summers (1989)). For example, the insertion can be into a gene such as the polyhedrin gene, by homologous double crossover recombination; insertion can also be into a restriction enzyme site engineered into the desired baculovirus gene. Miller et al., (1989), *Bioessays 4*:91. The DNA sequence, when cloned in place of the polyhedrin gene in the expression vector, is flanked both 5' and 3' by polyhedrin-specific sequences and is positioned downstream of the polyhedrin promoter.

The newly formed baculovirus expression vector is subsequently packaged into an infectious recombinant baculovirus. Homologous recombination occurs at low frequency (between about 1% and about 5%); thus, the majority of the virus produced after cotransfection is still wild-type virus. Therefore, a method is necessary to identify recombinant viruses. An advantage of the expression system is a visual screen allowing recombinant viruses to be distinguished. The polyhedrin protein, which is produced by the native virus, is produced at very high levels in the nuclei of infected cells at late times after viral infection. Accumulated polyhedrin protein forms occlusion bodies that also contain embedded particles. These occlusion bodies, up to 15 μm in size, are highly refractile, giving them a bright shiny appearance that is readily visualized under the light microscope. Cells infected with recombinant viruses lack occlusion bodies. To distinguish recombinant virus from wild-type virus, the transfection supernatant is plaqued onto a monolayer of insect cells by techniques known to those skilled in the art. Namely, the plaques are screened under the light microscope for the presence (indicative of wild-type virus) or absence (indicative of recombinant virus) of occlusion bodies. *Current Protocols in Microbiology* Vol. 2 (Ausubel et al. eds) at 16.8 (Supp. 10, 1990); Summers and Smith, *supra*; Miller et al. (1989).

Recombinant baculovirus expression vectors have been developed for infection into several insect cells. For example, recombinant baculoviruses have been developed for, *inter alia*: *Aedes aegypti* , *Autographa californica*, *Bombyx mori*, *Drosophila melanogaster*, *Spodoptera frugiperda*, and *Trichoplusia ni* (PCT Pub. No. WO 89/046699; Carbonell et al., (1985) *J. Virol. 56*:153; Wright (1986) *Nature 321*:718; Smith et al., (1983) *Mol. Cell. Biol. 3*:2156; and see generally, Fraser, *et al.* (1989) *In Vitro Cell. Dev. Biol. 25*:225).

Cells and cell culture media are commercially available for both direct and fusion expression of heterologous polypeptides in a baculovirus/expression system; cell culture technology is generally known to those skilled in the art. *See,* e.g., Summers and Smith *supra*.

The modified insect cells may then be grown in an appropriate nutrient medium, which allows for stable maintenance of the plasmid(s) present in the modified insect host. Where the expression product gene is under inducible control, the host may be grown to high density, and expression induced. Alternatively, where expression is constitutive, the product will be continuously expressed into the medium and the nutrient medium must be continuously circulated, while removing the product of interest and augmenting depleted nutrients. The product may be purified by such techniques as chromatography, e.g., HPLC, affinity chromatography, ion exchange chromatography, etc.; electrophoresis; density gradient centrifugation; solvent extraction, or the like. As appropriate, the product may be further purified, as required, so as to remove substantially any insect proteins which are also secreted in the medium or result from lysis of insect cells, so as to provide a product which is at least substantially free of host debris, e.g., proteins, lipids and polysaccharides.

In order to obtain protein expression, recombinant host cells derived from the transformants are incubated under conditions which allow expression of the recombinant protein encoding sequence. These conditions will vary, dependent upon the host cell selected. However, the conditions are readily ascertainable to those of ordinary skill in the art, based upon what is known in the art.

iv. Bacterial Systems

Bacterial expression techniques are known in the art. A bacterial promoter is any DNA sequence capable of binding bacterial RNA polymerase and initiating the downstream (3') transcription of a coding sequence (e.g. structural gene) into mRNA. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site and a transcription initiation site. A bacterial promoter may also have a second domain called an operator, that may overlap an adjacent RNA polymerase binding site at which RNA synthesis begins. The operator permits negative regulated (inducible) transcription, as a gene repressor protein may bind the operator and thereby inhibit transcription of a specific gene. Constitutive expression may occur in the absence of negative regulatory elements, such as the operator. In addition, positive regulation may be achieved by a gene activator protein binding sequence, which, if present is usually proximal (5') to the RNA polymerase binding sequence. An example of a gene activator protein is the catabolite activator protein (CAP), which helps

initiate transcription of the lac operon in Escherichia coli (*E. coli*) (Raibaud *et al.* (1984) *Annu. Rev. Genet. 18*:173). Regulated expression may therefore be either positive or negative, thereby either enhancing or reducing transcription.

Sequences encoding metabolic pathway enzymes provide particularly useful promoter sequences. Examples include promoter sequences derived from sugar metabolizing enzymes, such as galactose, lactose (*lac*) (Chang *et al.* (1977) *Nature 198*:1056), and maltose. Additional examples include promoter sequences derived from biosynthetic enzymes such as tryptophan (*trp*) (Goeddel *et al.* (1980) *Nuc. Acids Res. 8*:4057; Yelverton *et al.* (1981) *Nucl. Acids Res. 9*:731; U.S. Patent 4,738,921; EPO Publ. Nos. 036 776 and 121 775). The beta-lactamase (*bla*) promoter system (Weissmann (1981) "The cloning of interferon and other mistakes." In *Interferon 3* (ed. I. Gresser)), bacteriophage lambda PL (Shimatake *et al.* (1981) *Nature 292*:128) and T5 (U.S. Patent 4,689,406) promoter systems also provide useful promoter sequences.

In addition, synthetic promoters which do not occur in nature also function as bacterial promoters. For example, transcription activation sequences of one bacterial or bacteriophage promoter may be joined with the operon sequences of another bacterial or bacteriophage promoter, creating a synthetic hybrid promoter (U.S. Patent 4,551,433). For example, the *tac* promoter is a hybrid *trp-lac* promoter comprised of both *trp* promoter and *lac* operon sequences that is regulated by the *lac* repressor (Amann *et al.* (1983) *Gene 25*:167; de Boer *et al.* (1983) *Proc. Natl. Acad. Sci. 80*:21). Furthermore, a bacterial promoter can include naturally occurring promoters of non-bacterial origin that have the ability to bind bacterial RNA polymerase and initiate transcription. A naturally occurring promoter of non-bacterial origin can also be coupled with a compatible RNA polymerase to produce high levels of expression of some genes in prokaryotes. The bacteriophage T7 RNA polymerase/promoter system is an example of a coupled promoter system (Studier *et al.* (1986) *J. Mol. Biol. 189*:113; Tabor *et al.* (1985) *Proc Natl. Acad. Sci. 82*:1074). In addition, a hybrid promoter can also be comprised of a bacteriophage promoter and an *E. coli* operator region (EPO Publ. No. 267 851).

In addition to a functioning promoter sequence, an efficient ribosome binding site is also useful for the expression of foreign genes in prokaryotes. In *E. coli*, the ribosome binding site is called the Shine-Dalgarno (SD) sequence and includes an initiation codon (ATG) and a sequence 3-9 nucleotides in length located 3-11 nucleotides upstream of the initiation codon

(Shine *et al.* (1975) *Nature 254*:34). The SD sequence is thought to promote binding of mRNA to the ribosome by the pairing of bases between the SD sequence and the 3' end of *E. coli* 16S rRNA (Steitz *et al.* (1979) "Genetic signals and nucleotide sequences in messenger RNA." In *Biological Regulation and Development: Gene Expression* (ed. R.F. Goldberger)). To express eukaryotic genes and prokaryotic genes with weak ribosome-binding site, it is often necessary to optimize the distance between the SD sequence and the ATG of the eukaryotic gene (Sambrook *et al.* (1989) "Expression of cloned genes in Escherichia coli." In *Molecular Cloning: A Laboratory Manual*).

A DNA molecule may be expressed intracellularly. A promoter sequence may be directly linked with the DNA molecule, in which case the first amino acid at the N-terminus will always be a methionine, which is encoded by the ATG start codon. If desired, methionine at the N-terminus may be cleaved from the protein by *in vitro* incubation with cyanogen bromide or by either *in vivo* or *in vitro* incubation with a bacterial methionine N-terminal peptidase (EPO Publ. No. 219 237).

Fusion proteins provide an alternative to direct expression. Usually, a DNA sequence encoding the N-terminal portion of an endogenous bacterial protein, or other stable protein is fused to the 5' end of heterologous coding sequences. Upon expression, this construct will provide a fusion of the two amino acid sequences. For example, the bacteriophage lambda cell gene can be linked at the 5' terminus of a foreign gene and expressed in bacteria. The resulting fusion protein preferably retains a site for a processing enzyme (factor Xa) to cleave the bacteriophage protein from the foreign gene (Nagai *et al.* (1984) *Nature 309*:810). Fusion proteins can also be made with sequences from the *lac*Z (Jia *et al.* (1987) *Gene 60*:197), *trpE* (Allen et al. (1987) J. Biotechnol. 5:93; Makoff *et al.* (1989) *J. Gen. Microbiol. 135*:11), and *Chey* (EPO Publ. No. 324 647) genes. The DNA sequence at the junction of the two amino acid sequences may or may not encode a cleavable site. Another example is a ubiquitin fusion protein. Such a fusion protein is made with the ubiquitin region that preferably retains a site for a processing enzyme (e.g. ubiquitin specific processing-protease) to cleave the ubiquitin from the foreign protein. Through this method, native foreign protein can be isolated (Miller *et al.* (1989) *Bio/Technology 7*:698).

Alternatively, foreign proteins can also be secreted from the cell by creating chimeric DNA molecules that encode a fusion protein comprised of a signal peptide sequence fragment that provides for secretion of the foreign protein in bacteria (U.S. Patent 4,336,336). The

signal sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell. The protein is either secreted into the growth media (gram-positive bacteria) or into the periplasmic space, located between the inner and outer membrane of the cell (gram-negative bacteria). Preferably there are processing sites, which can be cleaved either *in vivo* or *in vitro* encoded between the signal peptide fragment and the foreign gene.

DNA encoding suitable signal sequences can be derived from genes for secreted bacterial proteins, such as the *E. coli* outer membrane protein gene (*ompA*) (Masui *et al.* (1983), in: *Experimental Manipulation of Gene Expression*; Ghrayeb *et al.* (1984) *EMBO J.* *3*:2437) and the *E. coli* alkaline phosphatase signal sequence (*phoA*) (Oka *et al.* (1985) *Proc. Natl. Acad. Sci. 82*:7212). As an additional example, the signal sequence of the alpha-amylase gene from various Bacillus strains can be used to secrete heterologous proteins from *B. subtilis* (Palva *et al.* (1982) *Proc. Natl. Acad. Sci. USA 79*:5582; EPO Publ. No. 244 042).

Usually, transcription termination sequences recognized by bacteria are regulatory regions located 3' to the translation stop codon, and thus together with the promoter flank the coding sequence. These sequences direct the transcription of an mRNA which can be translated into the polypeptide encoded by the DNA. Transcription termination sequences frequently include DNA sequences of about 50 nucleotides capable of forming stem loop structures that aid in terminating transcription. Examples include transcription termination sequences derived from genes with strong promoters, such as the *trp* gene in *E. coli* as well as other biosynthetic genes.

Usually, the above described components, comprising a promoter, signal sequence (if desired), coding sequence of interest, and transcription termination sequence, are put together into expression constructs. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as bacteria. The replicon will have a replication system, thus allowing it to be maintained in a prokaryotic host either for expression or for cloning and amplification. In addition, a replicon may be either a high or low copy number plasmid. A high copy number plasmid will generally have a copy number ranging from about 5 to about 200, and usually about 10 to about 150. A host containing a high copy number plasmid will preferably contain at least about 10, and more preferably at least about 20 plasmids. Either a high or low copy number

vector may be selected, depending upon the effect of the vector and the foreign protein on the host.

Alternatively, the expression constructs can be integrated into the bacterial genome with an integrating vector. Integrating vectors usually contain at least one sequence homologous to the bacterial chromosome that allows the vector to integrate. Integrations appear to result from recombinations between homologous DNA in the vector and the bacterial chromosome. For example, integrating vectors constructed with DNA from various Bacillus strains integrate into the Bacillus chromosome (EPO Publ. No. 127 328). Integrating vectors may also be comprised of bacteriophage or transposon sequences.

Usually, extrachromosomal and integrating expression constructs may contain selectable markers to allow for the selection of bacterial strains that have been transformed. Selectable markers can be expressed in the bacterial host and may include genes which render bacteria resistant to drugs such as ampicillin, chloramphenicol, erythromycin, kanamycin (neomycin), and tetracycline (Davies *et al.* (1978) *Annu. Rev. Microbiol. 32*:469). Selectable markers may also include biosynthetic genes, such as those in the histidine, tryptophan, and leucine biosynthetic pathways.

Alternatively, some of the above described components can be put together in transformation vectors. Transformation vectors are usually comprised of a selectable market that is either maintained in a replicon or developed into an integrating vector, as described above.

Expression and transformation vectors, either extra-chromosomal replicons or integrating vectors, have been developed for transformation into many bacteria. For example, expression vectors have been developed for, *inter alia*, the following bacteria: Bacillus subtilis (Palva *et al.* (1982) *Proc. Natl. Acad. Sci. USA 79*:5582; EPO Publ. Nos. 036 259 and 063 953; PCT Publ. No. WO 84/04541), Escherichia coli (Shimatake *et al.* (1981) *Nature 292*:128; Amann *et al.* (1985) *Gene 40*:183; Studier *et al.* (1986) *J. Mol. Biol. 189*:113; EPO Publ. Nos. 036 776, 136 829 and 136 907), Streptococcus cremoris (Powell *et al.* (1988) *Appl. Environ. Microbiol. 54*:655); Streptococcus lividans (Powell *et al.* (1988) *Appl. Environ. Microbiol. 54*:655), Streptomyces lividans (U.S. Patent 4,745,056).

Methods of introducing exogenous DNA into bacterial hosts are well-known in the art, and usually include either the transformation of bacteria treated with $CaCl_2$ or other agents, such as divalent cations and DMSO. DNA can also be introduced into bacterial cells by

electroporation. Transformation procedures usually vary with the bacterial species to be transformed. (See e.g., use of *Bacillus*: Masson *et al.* (1989) *FEMS Microbiol. Lett. 60*:273; Palva *et al.* (1982) *Proc. Natl. Acad. Sci. USA 79*:5582; EPO Publ. Nos. 036 259 and 063 953; PCT Publ. No. WO 84/04541; use of *Campylobacter*: Miller *et al.* (1988) *Proc. Natl. Acad. Sci. 85*:856; and Wang *et al.* (1990) *J. Bacteriol. 172*:949; use of *Escherichia coli:* Cohen *et al.* (1973) *Proc. Natl. Acad. Sci. 69*:2110; Dower *et al.* (1988) *Nucleic Acids Res. 16*:6127; Kushner (1978) "An improved method for transformation of *Escherichia coli* with ColE1-derived plasmids. In *Genetic Engineering: Proceedings of the International Symposium on Genetic Engineering* (eds. H.W. Boyer and S. Nicosia); Mandel *et al.* (1970) *J. Mol. Biol. 53*:159; Taketo (1988) *Biochim. Biophys. Acta 949*:318; use of *Lactobacillus*: Chassy *et al.* (1987) *FEMS Microbiol. Lett. 44*:173; use of *Pseudomonas*: Fiedler *et al.* (1988) *Anal. Biochem 170*:38; use of *Staphylococcus*: Augustin *et al.* (1990) *FEMS Microbiol. Lett. 66*:203; use of *Streptococcus*: Barany *et al.* (1980) *J. Bacteriol. 144*:698; Harlander (1987) "Transformation of Streptococcus lactis by electroporation, in: *Streptococcal Genetics* (ed. J. Ferretti and R. Curtiss III); Perry *et al.* (1981) *Infect. Immun. 32*:1295; Powell *et al.* (1988) *Appl. Environ. Microbiol. 54*:655; Somkuti *et al.* (1987) *Proc. 4th Evr. Cong. Biotechnology 1*:412.

v. Yeast Expression

Yeast expression systems are also known to one of ordinary skill in the art. A yeast promoter is any DNA sequence capable of binding yeast RNA polymerase and initiating the downstream (3') transcription of a coding sequence (e.g. structural gene) into mRNA. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site (the "TATA Box") and a transcription initiation site. A yeast promoter may also have a second domain called an upstream activator sequence (UAS), which, if present, is usually distal to the structural gene. The UAS permits regulated (inducible) expression. Constitutive expression occurs in the absence of a UAS. Regulated expression may be either positive or negative, thereby either enhancing or reducing transcription.

Yeast is a fermenting organism with an active metabolic pathway, therefore sequences encoding enzymes in the metabolic pathway provide particularly useful promoter sequences.

Examples include alcohol dehydrogenase (ADH) (EPO Publ. No. 284 044), enolase, glucokinase, glucose-6-phosphate isomerase, glyceraldehyde-3-phosphate-dehydrogenase (GAP or GAPDH), hexokinase, phosphofructokinase, 3-phosphoglycerate mutase, and pyruvate kinase (PyK) (EPO Publ. No. 329 203). The yeast *PHO5* gene, encoding acid phosphatase, also provides useful promoter sequences (Myanohara *et al.* (1983) *Proc. Natl. Acad. Sci. USA 80*:1).

In addition, synthetic promoters which do not occur in nature also function as yeast promoters. For example, UAS sequences of one yeast promoter may be joined with the transcription activation region of another yeast promoter, creating a synthetic hybrid promoter. Examples of such hybrid promoters include the ADH regulatory sequence linked to the GAP transcription activation region (U.S. Patent Nos. 4,876,197 and 4,880,734). Other examples of hybrid promoters include promoters which consist of the regulatory sequences of either the *ADH2*, *GAL4*, *GAL10*, OR *PHO5* genes, combined with the transcriptional activation region of a glycolytic enzyme gene such as GAP or PyK (EPO Publ. No. 164 556). Furthermore, a yeast promoter can include naturally occurring promoters of non-yeast origin that have the ability to bind yeast RNA polymerase and initiate transcription. Examples of such promoters include, *inter alia*, (Cohen *et al.* (1980) *Proc. Natl. Acad. Sci. USA 77*:1078; Henikoff *et al.* (1981) *Nature 283*:835; Hollenberg *et al.* (1981) *Curr. Topics Microbiol. Immunol. 96*:119; Hollenberg *et al.* (1979) "The Expression of Bacterial Antibiotic Resistance Genes in the Yeast Saccharomyces cerevisiae," in: *Plasmids of Medical, Environmental and Commercial Importance* (eds. K.N. Timmis and A. Puhler); Mercerau-Puigalon *et al.* (1980) *Gene 11*:163; Panthier *et al.* (1980) *Curr. Genet. 2*:109;).

A DNA molecule may be expressed intracellularly in yeast. A promoter sequence may be directly linked with the DNA molecule, in which case the first amino acid at the N-terminus of the recombinant protein will always be a methionine, which is encoded by the ATG start codon. If desired, methionine at the N-terminus may be cleaved from the protein by *in vitro* incubation with cyanogen bromide.

Fusion proteins provide an alternative for yeast expression systems, as well as in mammalian, plant, baculovirus, and bacterial expression systems. Usually, a DNA sequence encoding the N-terminal portion of an endogenous yeast protein, or other stable protein, is fused to the 5' end of heterologous coding sequences. Upon expression, this construct will provide a fusion of the two amino acid sequences. For example, the yeast or human

superoxide dismutase (SOD) gene, can be linked at the 5' terminus of a foreign gene and expressed in yeast. The DNA sequence at the junction of the two amino acid sequences may or may not encode a cleavable site. See e.g., EPO Publ. No. 196056. Another example is a ubiquitin fusion protein. Such a fusion protein is made with the ubiquitin region that preferably retains a site for a processing enzyme (e.g. ubiquitin-specific processing protease) to cleave the ubiquitin from the foreign protein. Through this method, therefore, native foreign protein can be isolated (e.g., WO88/024066).

Alternatively, foreign proteins can also be secreted from the cell into the growth media by creating chimeric DNA molecules that encode a fusion protein comprised of a leader sequence fragment that provide for secretion in yeast of the foreign protein. Preferably, there are processing sites encoded between the leader fragment and the foreign gene that can be cleaved either *in vivo* or *in vitro*. The leader sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell.

DNA encoding suitable signal sequences can be derived from genes for secreted yeast proteins, such as the yeast invertase gene (EPO Publ. No. 012 873; JPO Publ. No. 62:096,086) and the A-factor gene (U.S. Patent 4,588,684). Alternatively, leaders of non-yeast origin, such as an interferon leader, exist that also provide for secretion in yeast (EPO Publ. No. 060 057).

A preferred class of secretion leaders are those that employ a fragment of the yeast alpha-factor gene, which contains both a "pre" signal sequence, and a "pro" region. The types of alpha-factor fragments that can be employed include the full-length pre-pro alpha factor leader (about 83 amino acid residues) as well as truncated alpha-factor leaders (usually about 25 to about 50 amino acid residues) (U.S. Patent Nos. 4,546,083 and 4,870,008; EPO Publ. No. 324 274). Additional leaders employing an alpha-factor leader fragment that provides for secretion include hybrid alpha-factor leaders made with a presequence of a first yeast, but a pro-region from a second yeast alphafactor. (See e.g., PCT Publ. No. WO 89/02463.)

Usually, transcription termination sequences recognized by yeast are regulatory regions located 3' to the translation stop codon, and thus together with the promoter flank the coding sequence. These sequences direct the transcription of an mRNA which can be translated into the polypeptide encoded by the DNA. Examples of transcription terminator

sequence and other yeast-recognized termination sequences, such as those coding for glycolytic enzymes.

Usually, the above described components, comprising a promoter, leader (if desired), coding sequence of interest, and transcription termination sequence, are put together into expression constructs. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as yeast or bacteria. The replicon may have two replication systems, thus allowing it to be maintained, for example, in yeast for expression and in a prokaryotic host for cloning and amplification. Examples of such yeast-bacteria shuttle vectors include YEp24 (Botstein *et al.* (1979) *Gene 8*:17-24), pCl/1 (Brake *et al.* (1984) *Proc. Natl. Acad. Sci USA 81*:4642-4646), and YRp17 (Stinchcomb *et al.* (1982) *J. Mol. Biol. 158*:157). In addition, a replicon may be either a high or low copy number plasmid. A high copy number plasmid will generally have a copy number ranging from about 5 to about 200, and usually about 10 to about 150. A host containing a high copy number plasmid will preferably have at least about 10, and more preferably at least about 20. Enter a high or low copy number vector may be selected, depending upon the effect of the vector and the foreign protein on the host. See e.g., Brake *et al.*, *supra.*

Alternatively, the expression constructs can be integrated into the yeast genome with an integrating vector. Integrating vectors usually contain at least one sequence homologous to a yeast chromosome that allows the vector to integrate, and preferably contain two homologous sequences flanking the expression construct. Integrations appear to result from recombinations between homologous DNA in the vector and the yeast chromosome (Orr-Weaver *et al.* (1983) *Methods in Enzymol. 101*:228-245). An integrating vector may be directed to a specific locus in yeast by selecting the appropriate homologous sequence for inclusion in the vector. See Orr-Weaver *et al.*, *supra.* One or more expression construct may integrate, possibly affecting levels of recombinant protein produced (Rine *et al.* (1983) *Proc. Natl. Acad. Sci. USA 80*:6750). The chromosomal sequences included in the vector can occur either as a single segment in the vector, which results in the integration of the entire vector, or two segments homologous to adjacent segments in the chromosome and flanking the expression construct in the vector, which can result in the stable integration of only the expression construct.

Usually, extrachromosomal and integrating expression constructs may contain selectable markers to allow for the selection of yeast strains that have been transformed. Selectable markers may include biosynthetic genes that can be expressed in the yeast host, such as *ADE2*, *HIS4*, *LEU2*, *TRP1*, and *ALG7*, and the G418 resistance gene, which confer resistance in yeast cells to tunicamycin and G418, respectively. In addition, a suitable selectable marker may also provide yeast with the ability to grow in the presence of toxic compounds, such as metal. For example, the presence of *CUP1* allows yeast to grow in the presence of copper ions (Butt *et al.* (1987) *Microbiol, Rev. 51*:351).

Alternatively, some of the above described components can be put together into transformation vectors. Transformation vectors are usually comprised of a selectable marker that is either maintained in a replicon or developed into an integrating vector, as described above.

Expression and transformation vectors, either extrachromosomal replicons or integrating vectors, have been developed for transformation into many yeasts. For example, expression vectors and methods of introducing exogenous DNA into yeast hosts have been developed for, *inter alia*, the following yeasts: *Candida albicans* (Kurtz, *et al.* (1986) *Mol. Cell. Biol. 6*:142); *Candida maltosa* (Kunze, *et al.* (1985) *J. Basic Microbiol. 25*:141); *Hansenula polymorpha* (Gleeson, *et al.* (1986) *J. Gen. Microbiol. 132*:3459; Roggenkamp *et al.* (1986) *Mol. Gen. Genet. 202*:302); *Kluyveromyces fragilis* (Das, *et al.* (1984) *J. Bacteriol. 158*:1165); *Kluyveromyces lactis* (De Louvencourt *et al.* (1983) *J. Bacteriol. 154*:737; Van den Berg *et al.* (1990) *Bio/Technology 8*:135); *Pichia guillerimondii* (Kunze *et al.* (1985) *J. Basic Microbiol. 25*:141); *Pichia pastoris* (Cregg, *et al.* (1985) *Mol. Cell. Biol. 5*:3376; U.S. Patent Nos. 4,837,148 and 4,929,555); *Saccharomyces cerevisiae* (Hinnen *et al.* (1978) *Proc. Natl. Acad. Sci. USA 75*:1929; Ito *et al.* (1983) *J. Bacteriol. 153*:163); *Schizosaccharomyces pombe* (Beach and Nurse (1981) *Nature 300*:706); and *Yarrowia lipolytica* (Davidow, *et al.* (1985) *Curr. Genet. 10*:380471 Gaillardin, *et al.* (1985) *Curr. Genet. 10*:49).

Methods of introducing exogenous DNA into yeast hosts are well-known in the art, and usually include either the transformation of spheroplasts or of intact yeast cells treated with alkali cations. Transformation procedures usually vary with the yeast species to be transformed. See e.g., [Kurtz *et al.* (1986) *Mol. Cell. Biol. 6*:142; Kunze *et al.* (1985) *J. Basic Microbiol. 25*:141; Candida]; [Gleeson *et al.* (1986) *J. Gen. Microbiol. 132*:3459; Roggenkamp *et al.* (1986) *Mol. Gen. Genet. 202*:302; Hansenula]; [Das *et al.* (1984) *J.*

Bacteriol. 158:1165; De Louvencourt *et al.* (1983) *J. Bacteriol. 154*:1165; Van den Berg *et al.* (1990) *Bio/Technology 8*:135; Kluyveromyces]; [Cregg *et al.* (1985) *Mol. Cell. Biol. 5*:3376; Kunze *et al.* (1985) *J. Basic Microbiol. 25*:141; U.S. Patent Nos. 4,837,148 and 4,929,555; Pichia]; [Hinnen *et al.* (1978) *Proc. Natl. Acad. Sci. USA 75*;1929; Ito *et al.* (1983) *J. Bacteriol. 153*:163 Saccharomyces]; [Beach and Nurse (1981) *Nature 300*:706; Schizosaccharomyces]; [Davidow *et al.* (1985) *Curr. Genet. 10*:39; Gaillardin *et al.* (1985) *Curr. Genet. 10*:49; Yarrowia].

Definitions

A composition containing X is "substantially free of" Y when at least 85% by weight of the total X+Y in the composition is X. Preferably, X comprises at least about 90% by weight of the total of X+Y in the composition, more preferably at least about 95% or even 99% by weight.

A "conserved" *Neisseria* amino acid fragment or protein is one that is present in a particular Neisserial protein in at least x% of *Neisseria*. The value of x may be 50% or more, e.g., 66%, 75%, 80%, 90%, 95% or even 100% (i.e. the amino acid is found in the protein in question in all *Neisseria*). In order to determine whether an animo acid is "conserved" in a particular Neisserial protein, it is necessary to compare that amino acid residue in the sequences of the protein in question from a plurality of different *Neisseria* (a reference population). The reference population may include a number of different *Neisseria* species or may include a single species. The reference population may include a number of different serogroups of a particular species or a single serogroup. A preferred reference population consists of the 5 most common *Neisseria* strains.

The term "heterologous" refers to two biological components that are not found together in nature. The components may be host cells, genes, or regulatory regions, such as promoters. Although the heterologous components are not found together in nature, they can function together, as when a promoter heterologous to a gene is operably linked to the gene. Another example is where a Neisserial sequence is heterologous to a mouse host cell.

"Epitope" means antigenic determinant, and may elicit a cellular and/or humoral response.

Conditions for "high stringency" are 65 degrees C in 0.1 xSSC 0.5% SDS solution.

An "origin of replication" is a polynucleotide sequence that initiates and regulates replication of polynucleotides, such as an expression vector. The origin of replication behaves as an autonomous unit of polynucleotide replication within a cell, capable of replication under its own control. An origin of replication may be needed for a vector to replicate in a particular host cell. With certain origins of replication, an expression vector can be reproduced at a high copy number in the presence of the appropriate proteins within the cell. Examples of origins are the autonomously replicating sequences, which are effective in yeast; and the viral T-antigen, effective in COS-7 cells.

A "mutant" sequence is defined as a DNA, RNA or amino acid sequence differing from but having homology with the native or disclosed sequence. Depending on the particular sequence, the degree of homology (sequence identity) between the native or disclosed sequence and the mutant sequence is preferably greater than 50% (e.g., 60%, 70%, 80%, 90%, 95%, 99% or more) which is calculated as described above. As used herein, an "allelic variant" of a nucleic acid molecule, or region, for which nucleic acid sequence is provided herein is a nucleic acid molecule, or region, that occurs at essentially the same locus in the genome of another or second isolate, and that, due to natural variation caused by, for example, mutation or recombination, has a similar but not identical nucleic acid sequence. A coding region allelic variant typically encodes a protein having similar activity to that of the protein encoded by the gene to which it is being compared. An allelic variant can also comprise an alteration in the 5' or 3' untranslated regions of the gene, such as in regulatory control regions. (see, for example, U.S. Patent 5,753,235).

Antibodies

As used herein, the term "antibody" refers to a polypeptide or group of polypeptides composed of at least one antibody combining site. An "antibody combining site" is the three-dimensional binding space with an internal surface shape and charge distribution complementary to the features of an epitope of an antigen, which allows a binding of the antibody with the antigen. "Antibody" includes, for example, vertebrate antibodies, hybrid antibodies, chimeric antibodies, humanized antibodies, altered antibodies, univalent antibodies, Fab proteins, and single domain antibodies.

Antibodies against the proteins of the invention are useful for affinity chromatography, immunoassays, and distinguishing/identifying *Neisseria* menB proteins.

Antibodies elicited against the proteins of the present invention bind to antigenic polypeptides or proteins or protein fragments that are present and specifically associated with strains of *Neisseria meningitidis* menB. In some instances, these antigens may be associated with specific strains, such as those antigens specific for the menB strains. The antibodies of the invention may be immobilized to a matrix and utilized in an immunoassay or on an affinity chromatography column, to enable the detection and/or separation of polypeptides, proteins or protein fragments or cells comprising such polypeptides, proteins or protein fragments. Alternatively, such polypeptides, proteins or protein fragments may be immobilized so as to detect antibodies bindably specific thereto.

Antibodies to the proteins of the invention, both polyclonal and monoclonal, may be prepared by conventional methods. In general, the protein is first used to immunize a suitable animal, preferably a mouse, rat, rabbit or goat. Rabbits and goats are preferred for the preparation of polyclonal sera due to the volume of serum obtainable, and the availability of labeled anti-rabbit and anti-goat antibodies. Immunization is generally performed by mixing or emulsifying the protein in saline, preferably in an adjuvant such as Freund's complete adjuvant, and injecting the mixture or emulsion parenterally (generally subcutaneously or intramuscularly). A dose of 50-200 μg/injection is typically sufficient. Immunization is generally boosted 2-6 weeks later with one or more injections of the protein in saline, preferably using Freund's incomplete adjuvant. One may alternatively generate antibodies by in vitro immunization using methods known in the art, which for the purposes of this invention is considered equivalent to *in vivo* immunization. Polyclonal antisera is obtained by bleeding the immunized animal into a glass or plastic container, incubating the blood at 25°C for one hour, followed by incubating at 4°C for 2-18 hours. The serum is recovered by centrifugation (e.g., 1,000*g* for 10 minutes). About 20-50 ml per bleed may be obtained from rabbits.

Monoclonal antibodies are prepared using the standard method of Kohler & Milstein (*Nature* (1975) 256:495-96), or a modification thereof. Typically, a mouse or rat is immunized as described above. However, rather than bleeding the animal to extract serum, the spleen (and optionally several large lymph nodes) is removed and dissociated into single cells. If desired, the spleen cells may be screened (after removal of nonspecifically adherent cells) by applying a cell suspension to a plate or well coated with the protein antigen. B-cells that express membrane-bound immunoglobulin specific for the antigen bind to the plate, and

are not rinsed away with the rest of the suspension. Resulting B-cells, or all dissociated spleen cells, are then induced to fuse with myeloma cells to form hybridomas, and are cultured in a selective medium (e.g., hypoxanthine, aminopterin, thymidine medium, "HAT"). The resulting hybridomas are plated by limiting dilution, and are assayed for the production of antibodies which bind specifically to the immunizing antigen (and which do not bind to unrelated antigens). The selected MAb-secreting hybridomas are then cultured either *in vitro* (e.g., in tissue culture bottles or hollow fiber reactors), or *in vivo* (as ascites in mice).

If desired, the antibodies (whether polyclonal or monoclonal) may be labeled using conventional techniques. Suitable labels include fluorophores, chromophores, radioactive atoms (particularly ^{32}P and ^{125}I), electron-dense reagents, enzymes, and ligands having specific binding partners. Enzymes are typically detected by their activity. For example, horseradish peroxidase is usually detected by its ability to convert 3,3',5,5'-tetramethylbenzidine (TMB) to a blue pigment, quantifiable with a spectrophotometer. "Specific binding partner" refers to a protein capable of binding a ligand molecule with high specificity, as for example in the case of an antigen and a monoclonal antibody specific therefor. Other specific binding partners include biotin and avidin or streptavidin, IgG and protein A, and the numerous receptor-ligand couples known in the art. It should be understood that the above description is not meant to categorize the various labels into distinct classes, as the same label may serve in several different modes. For example, ^{125}I may serve as a radioactive label or as an electron-dense reagent. HRP may serve as enzyme or as antigen for a MAb. Further, one may combine various labels for desired effect. For example, MAbs and avidin also require labels in the practice of this invention: thus, one might label a MAb with biotin, and detect its presence with avidin labeled with ^{125}I, or with an anti-biotin MAb labeled with HRP. Other permutations and possibilities will be readily apparent to those of ordinary skill in the art, and are considered as equivalents within the scope of the instant invention.

Antigens, immunogens, polypeptides, proteins or protein fragments of the present invention elicit formation of specific binding partner antibodies. These antigens, immunogens, polypeptides, proteins or protein fragments of the present invention comprise immunogenic compositions of the present invention. Such immunogenic compositions may further comprise or include adjuvants, carriers, or other compositions that promote or enhance

or stabilize the antigens, polypeptides, proteins or protein fragments of the present invention. Such adjuvants and carriers will be readily apparent to those of ordinary skill in the art.

Pharmaceutical Compositions

Pharmaceutical compositions can comprise (include) either polypeptides, antibodies, or nucleic acid of the invention. The pharmaceutical compositions will comprise a therapeutically effective amount of either polypeptides, antibodies, or polynucleotides of the claimed invention.

The term "therapeutically effective amount" as used herein refers to an amount of a therapeutic agent to treat, ameliorate, or prevent a desired disease or condition, or to exhibit a detectable therapeutic or preventative effect. The effect can be detected by, for example, chemical markers or antigen levels. Therapeutic effects also include reduction in physical symptoms, such as decreased body temperature, when given to a patient that is febrile. The precise effective amount for a subject will depend upon the subject's size and health, the nature and extent of the condition, and the therapeutics or combination of therapeutics selected for administration. Thus, it is not useful to specify an exact effective amount in advance. However, the effective amount for a given situation can be determined by routine experimentation and is within the judgment of the clinician.

For purposes of the present invention, an effective dose will be from about 0.01 mg/kg to 50 mg/kg or 0.05 mg/kg to about 10 mg/kg of the DNA constructs in the individual to which it is administered.

A pharmaceutical composition can also contain a pharmaceutically acceptable carrier. The term "pharmaceutically acceptable carrier" refers to a carrier for administration of a therapeutic agent, such as antibodies or a polypeptide, genes, and other therapeutic agents. The term refers to any pharmaceutical carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition, and which may be administered without undue toxicity. Suitable carriers may be large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, and inactive virus particles. Such carriers are well known to those of ordinary skill in the art.

Pharmaceutically acceptable salts can be used therein, for example, mineral acid salts such as hydrochlorides, hydrobromides, phosphates, sulfates, and the like; and the salts of

organic acids such as acetates, propionates, malonates, benzoates, and the like. A thorough discussion of pharmaceutically acceptable excipients is available in Remington's Pharmaceutical Sciences (Mack Pub. Co., N.J. 1991).

Pharmaceutically acceptable carriers in therapeutic compositions may contain liquids such as water, saline, glycerol and ethanol. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles. Typically, the therapeutic compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared. Liposomes are included within the definition of a pharmaceutically acceptable carrier.

Delivery Methods

Once formulated, the compositions of the invention can be administered directly to the subject. The subjects to be treated can be animals; in particular, human subjects can be treated.

Direct delivery of the compositions will generally be accomplished by injection, either subcutaneously, intraperitoneally, intravenously or intramuscularly or delivered to the interstitial space of a tissue. The compositions can also be administered into a lesion. Other modes of administration include oral and pulmonary administration, suppositories, and transdermal and transcutaneous applications, needles, and gene guns or hyposprays. Dosage treatment may be a single dose schedule or a multiple dose schedule.

Vaccines

Vaccines according to the invention may either be prophylactic (i.e., to prevent infection) or therapeutic (i.e., to treat disease after infection).

Such vaccines comprise immunizing antigen(s) or immunogen(s), immunogenic polypeptide, protein(s) or protein fragments, or nucleic acids (e.g., ribonucleic acid or deoxyribonucleic acid), usually in combination with "pharmaceutically acceptable carriers," which include any carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition. Suitable carriers are typically large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, lipid aggregates (such as oil droplets or

liposomes), and inactive virus particles. Such carriers are well known to those of ordinary skill in the art. Additionally, these carriers may function as immunostimulating agents ("adjuvants"). Furthermore, the immunogen or antigen may be conjugated to a bacterial toxoid, such as a toxoid from diphtheria, tetanus, cholera, *H. pylori*, etc. pathogens.

Preferred adjuvants to enhance effectiveness of the composition include, but are not limited to: (1) aluminum salts (alum), such as aluminum hydroxide, aluminum phosphate, aluminum sulfate, etc; (2) oil-in-water emulsion formulations (with or without other specific immunostimulating agents such as muramyl peptides (see below) or bacterial cell wall components), such as for example (a) MF59 (PCT Publ. No. WO 90/14837), containing 5% Squalene, 0.5% Tween 80, and 0.5% Span 85 (optionally containing various amounts of MTP-PE (see below), although not required) formulated into submicron particles using a microfluidizer such as Model 110Y microfluidizer (Microfluidics, Newton, MA), (b) SAF, containing 10% Squalane, 0.4% Tween 80, 5% pluronic-blocked polymer L121, and thr-MDP (see below) either microfluidized into a submicron emulsion or vortexed to generate a larger particle size emulsion, and (c) Ribi™ adjuvant system (RAS), (Ribi Immunochem, Hamilton, MT) containing 2% Squalene, 0.2% Tween 80, and one or more bacterial cell wall components from the group consisting of monophosphorylipid A (MPL), trehalose dimycolate (TDM), and cell wall skeleton (CWS), preferably MPL + CWS (Detox™);
(3) saponin adjuvants, such as Stimulon™ (Cambridge Bioscience, Worcester, MA) may be used or particles generated therefrom such as ISCOMs (immunostimulating complexes);
(4) Complete Freund's Adjuvant (CFA) and Incomplete Freund's Adjuvant (IFA);
(5) cytokines, such as interleukins (e.g., IL-1, IL-2, IL-4, IL-5, IL-6, IL-7, IL-12, *etc.*), interferons (e.g., gamma interferon), macrophage colony stimulating factor (M-CSF), tumor necrosis factor (TNF), etc; (6) detoxified mutants of a bacterial ADP-ribosylating toxin such as a cholera toxin (CT), a pertussis toxin (PT), or an E. coli heat-labile toxin (LT), particularly LT-K63, LT-R72, CT-S109, PT-K9/G129; see, e.g., WO 93/13302 and WO 92/19265; and
(7) other substances that act as immunostimulating agents to enhance the effectiveness of the composition. Alum and MF59 are preferred.

As mentioned above, muramyl peptides include, but are not limited to, N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acetyl-normuramyl-L-alanyl-D-isoglutamine (nor-MDP), N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-*sn*-glycero-3-huydroxyphosphoryloxy)-ethylamine (MTP-PE), *etc.*

The vaccine compositions comprising immunogenic compositions (e.g., which may include the antigen, pharmaceutically acceptable carrier, and adjuvant) typically will contain diluents, such as water, saline, glycerol, ethanol, etc. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles. Alternatively, vaccine compositions comprising immunogenic compositions may comprise an antigen, polypeptide, protein, protein fragment or nucleic acid in a pharmaceutically acceptable carrier.

More specifically, vaccines comprising immunogenic compositions comprise an immunologically effective amount of the immunogenic polypeptides, as well as any other of the above-mentioned components, as needed. By "immunologically effective amount", it is meant that the administration of that amount to an individual, either in a single dose or as part of a series, is effective for treatment or prevention. This amount varies depending upon the health and physical condition of the individual to be treated, the taxonomic group of individual to be treated (e.g., nonhuman primate, primate, *etc.*), the capacity of the individual's immune system to synthesize antibodies, the degree of protection desired, the formulation of the vaccine, the treating doctor's assessment of the medical situation, and other relevant factors. It is expected that the amount will fall in a relatively broad range that can be determined through routine trials.

Typically, the vaccine compositions or immunogenic compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared. The preparation also may be emulsified or encapsulated in liposomes for enhanced adjuvant effect, as discussed above under pharmaceutically acceptable carriers.

The immunogenic compositions are conventionally administered parenterally, e.g., by injection, either subcutaneously or intramuscularly. Additional formulations suitable for other modes of administration include oral and pulmonary formulations, suppositories, and transdermal and transcutaneous applications. Dosage treatment may be a single dose schedule or a multiple dose schedule. The vaccine may be administered in conjunction with other immunoregulatory agents.

As an alternative to protein-based vaccines, DNA vaccination may be employed (e.g., Robinson & Torres (1997) *Seminars in Immunology* 9:271-283; Donnelly *et al.* (1997) *Annu Rev Immunol* 15:617-648).

Gene Delivery Vehicles

Gene therapy vehicles for delivery of constructs, including a coding sequence of a therapeutic of the invention, to be delivered to the mammal for expression in the mammal, can be administered either locally or systemically. These constructs can utilize viral or non-viral vector approaches in *in vivo* or *ex vivo* modality. Expression of such coding sequence can be induced using endogenous mammalian or heterologous promoters. Expression of the coding sequence in vivo can be either constitutive or regulated.

The invention includes gene delivery vehicles capable of expressing the contemplated nucleic acid sequences. The gene delivery vehicle is preferably a viral vector and, more preferably, a retroviral, adenoviral, adeno-associated viral (AAV), herpes viral, or alphavirus vector. The viral vector can also be an astrovirus, coronavirus, orthomyxovirus, papovavirus, paramyxovirus, parvovirus, picornavirus, poxvirus, or togavirus viral vector. See generally, Jolly (1994) *Cancer Gene Therapy* 1:51-64; Kimura (1994) *Human Gene Therapy* 5:845-852; Connelly (1995) *Human Gene Therapy* 6:185-193; and Kaplitt (1994) *Nature Genetics* 6:148-153.

Retroviral vectors are well known in the art, including B, C and D type retroviruses, xenotropic retroviruses (for example, NZB-X1, NZB-X2 and NZB9-1 (see O'Neill (1985) *J. Virol.* 53:160) polytropic retroviruses e.g., MCF and MCF-MLV (see Kelly (1983) *J. Virol.* 45:291), spumaviruses and lentiviruses. See RNA Tumor Viruses, Second Edition, Cold Spring Harbor Laboratory, 1985.

Portions of the retroviral gene therapy vector may be derived from different retroviruses. For example, retrovector LTRs may be derived from a Murine Sarcoma Virus, a tRNA binding site from a Rous Sarcoma Virus, a packaging signal from a Murine Leukemia Virus, and an origin of second strand synthesis from an Avian Leukosis Virus.

These recombinant retroviral vectors may be used to generate transduction competent retroviral vector particles by introducing them into appropriate packaging cell lines (see US patent 5,591,624). Retrovirus vectors can be constructed for site-specific integration into host cell DNA by incorporation of a chimeric integrase enzyme into the retroviral particle (see WO96/37626). It is preferable that the recombinant viral vector is a replication defective recombinant virus.

Packaging cell lines suitable for use with the above-described retrovirus vectors are well known in the art, are readily prepared (see WO95/30763 and WO92/05266), and can be used to create producer cell lines (also termed vector cell lines or "VCLs") for the production of recombinant vector particles. Preferably, the packaging cell lines are made from human parent cells (e.g., HT1080 cells) or mink parent cell lines, which eliminates inactivation in human serum.

Preferred retroviruses for the construction of retroviral gene therapy vectors include Avian Leukosis Virus, Bovine Leukemia, Virus, Murine Leukemia Virus, Mink-Cell Focus-Inducing Virus, Murine Sarcoma Virus, Reticuloendotheliosis Virus and Rous Sarcoma Virus. Particularly preferred Murine Leukemia Viruses include 4070A and 1504A (Hartley and Rowe (1976) *J Virol* 19:19-25), Abelson (ATCC No. VR-999), Friend (ATCC No. VR-245), Graffi, Gross (ATCC Nol VR-590), Kirsten, Harvey Sarcoma Virus and Rauscher (ATCC No. VR-998) and Moloney Murine Leukemia Virus (ATCC No. VR-190). Such retroviruses may be obtained from depositories or collections such as the American Type Culture Collection ("ATCC") in Rockville, Maryland or isolated from known sources using commonly available techniques.

Exemplary known retroviral gene therapy vectors employable in this invention include those described in patent applications GB2200651, EP0415731, EP0345242, EP0334301, WO89/02468; WO89/05349, WO89/09271, WO90/02806, WO90/07936, WO94/03622, WO93/25698, WO93/25234, WO93/11230, WO93/10218, WO91/02805, WO91/02825, WO95/07994, US 5,219,740, US 4,405,712, US 4,861,719, US 4,980,289, US 4,777,127, US 5,591,624. See also Vile (1993) *Cancer Res* 53:3860-3864; Vile (1993) *Cancer Res* 53:962-967; Ram (1993) *Cancer Res* 53 (1993) 83-88; Takamiya (1992) *J Neurosci Res* 33:493-503; Baba (1993) *J Neurosurg* 79:729-735; Mann (1983) *Cell* 33:153; Cane (1984) *Proc Natl Acad Sci* 81:6349; and Miller (1990) *Human Gene Therapy* 1.

Human adenoviral gene therapy vectors are also known in the art and employable in this invention. See, for example, Berkner (1988) *Biotechniques* 6:616 and Rosenfeld (1991) *Science* 252:431, and WO93/07283, WO93/06223, and WO93/07282. Exemplary known adenoviral gene therapy vectors employable in this invention include those described in the above referenced documents and in WO94/12649, WO93/03769, WO93/19191, WO94/28938, WO95/11984, WO95/00655, WO95/27071, WO95/29993, WO95/34671, WO96/05320, WO94/08026, WO94/11506, WO93/06223, WO94/24299, WO95/14102,

WO95/24297, WO95/02697, WO94/28152, WO94/24299, WO95/09241, WO95/25807, WO95/05835, WO94/18922 and WO95/09654. Alternatively, administration of DNA linked to killed adenovirus as described in Curiel (1992) *Hum. Gene Ther.* 3:147-154 may be employed. The gene delivery vehicles of the invention also include adenovirus associated virus (AAV) vectors. Leading and preferred examples of such vectors for use in this invention are the AAV-2 based vectors disclosed in Srivastava, WO93/09239. Most preferred AAV vectors comprise the two AAV inverted terminal repeats in which the native D-sequences are modified by substitution of nucleotides, such that at least 5 native nucleotides and up to 18 native nucleotides, preferably at least 10 native nucleotides up to 18 native nucleotides, most preferably 10 native nucleotides are retained and the remaining nucleotides of the D-sequence are deleted or replaced with non-native nucleotides. The native D-sequences of the AAV inverted terminal repeats are sequences of 20 consecutive nucleotides in each AAV inverted terminal repeat (i.e., there is one sequence at each end) which are not involved in HP formation. The non-native replacement nucleotide may be any nucleotide other than the nucleotide found in the native D-sequence in the same position. Other employable exemplary AAV vectors are pWP-19, pWN-1, both of which are disclosed in Nahreini (1993) *Gene* 124:257-262. Another example of such an AAV vector is psub201 (see Samulski (1987) *J. Virol.* 61:3096). Another exemplary AAV vector is the Double-D ITR vector. Construction of the Double-D ITR vector is disclosed in US Patent 5,478,745. Still other vectors are those disclosed in Carter US Patent 4,797,368 and Muzyczka US Patent 5,139,941, Chartejee US Patent 5,474,935, and Kotin WO94/288157. Yet a further example of an AAV vector employable in this invention is SSV9AFABTKneo, which contains the AFP enhancer and albumin promoter and directs expression predominantly in the liver. Its structure and construction are disclosed in Su (1996) *Human Gene Therapy* 7:463-470. Additional AAV gene therapy vectors are described in US 5,354,678, US 5,173,414, US 5,139,941, and US 5,252,479.

The gene therapy vectors comprising sequences of the invention also include herpes vectors. Leading and preferred examples are herpes simplex virus vectors containing a sequence encoding a thymidine kinase polypeptide such as those disclosed in US 5,288,641 and EP0176170 (Roizman). Additional exemplary herpes simplex virus vectors include HFEM/ICP6-LacZ disclosed in WO95/04139 (Wistar Institute), pHSVlac described in Geller (1988) *Science* 241:1667-1669 and in WO90/09441 and WO92/07945, HSV Us3::pgC-lacZ

described in Fink (1992) *Human Gene Therapy* 3:11-19 and HSV 7134, 2 RH 105 and GAL4 described in EP 0453242 (Breakefield), and those deposited with the ATCC as accession numbers ATCC VR-977 and ATCC VR-260.

Also contemplated are alpha virus gene therapy vectors that can be employed in this invention. Preferred alpha virus vectors are Sindbis viruses vectors. Togaviruses, Semliki Forest virus (ATCC VR-67; ATCC VR-1247), Middleberg virus (ATCC VR-370), Ross River virus (ATCC VR-373; ATCC VR-1246), Venezuelan equine encephalitis virus (ATCC VR923; ATCC VR-1250; ATCC VR-1249; ATCC VR-532), and those described in US patents 5,091,309, 5,217,879, and WO92/10578. More particularly, those alpha virus vectors described in U.S. Serial No. 08/405,627, filed March 15, 1995,WO94/21792, WO92/10578, WO95/07994, US 5,091,309 and US 5,217,879 are employable. Such alpha viruses may be obtained from depositories or collections such as the ATCC in Rockville, Maryland or isolated from known sources using commonly available techniques. Preferably, alphavirus vectors with reduced cytotoxicity are used (see USSN 08/679640).

DNA vector systems such as eukarytic layered expression systems are also useful for expressing the nucleic acids of the invention. SeeWO95/07994 for a detailed description of eukaryotic layered expression systems. Preferably, the eukaryotic layered expression systems of the invention are derived from alphavirus vectors and most preferably from Sindbis viral vectors.

Other viral vectors suitable for use in the present invention include those derived from poliovirus, for example ATCC VR-58 and those described in Evans, Nature 339 (1989) 385 and Sabin (1973) *J. Biol. Standardization* 1:115; rhinovirus, for example ATCC VR-1110 and those described in Arnold (1990) *J Cell Biochem* L401; pox viruses such as canary pox virus or vaccinia virus, for example ATCC VR-111 and ATCC VR-2010 and those described in Fisher-Hoch (1989) *Proc Natl Acad Sci* 86:317; Flexner (1989) *Ann NY Acad Sci* 569:86, Flexner (1990) *Vaccine* 8:17; in US 4,603,112 and US 4,769,330 and WO89/01973; SV40 virus, for example ATCC VR-305 and those described in Mulligan (1979) *Nature* 277:108 and Madzak (1992) *J Gen Virol* 73:1533; influenza virus, for example ATCC VR-797 and recombinant influenza viruses made employing reverse genetics techniques as described in US 5,166,057 and in Enami (1990) *Proc Natl Acad Sci* 87:3802-3805; Enami & Palese (1991) *J Virol* 65:2711-2713 and Luytjes (1989) *Cell* 59:110, (see also McMichael (1983) *NEJ Med* 309:13, and Yap (1978) *Nature* 273:238 and Nature (1979) 277:108); human

immunodeficiency virus as described in EP-0386882 and in Buchschacher (1992) *J. Virol.* 66:2731; measles virus, for example ATCC VR-67 and VR-1247 and those described in EP-0440219; Aura virus, for example ATCC VR-368; Bebaru virus, for example ATCC VR-600 and ATCC VR-1240; Cabassou virus, for example ATCC VR-922; Chikungunya virus, for example ATCC VR-64 and ATCC VR-1241; Fort Morgan Virus, for example ATCC VR-924; Getah virus, for example ATCC VR-369 and ATCC VR-1243; Kyzylagach virus, for example ATCC VR-927; Mayaro virus, for example ATCC VR-66; Mucambo virus, for example ATCC VR-580 and ATCC VR-1244; Ndumu virus, for example ATCC VR-371; Pixuna virus, for example ATCC VR-372 and ATCC VR-1245; Tonate virus, for example ATCC VR-925; Triniti virus, for example ATCC VR-469; Una virus, for example ATCC VR-374; Whataroa virus, for example ATCC VR-926; Y-62-33 virus, for example ATCC VR-375; O'Nyong virus, Eastern encephalitis virus, for example ATCC VR-65 and ATCC VR-1242; Western encephalitis virus, for example ATCC VR-70, ATCC VR-1251, ATCC VR-622 and ATCC VR-1252; and coronavirus, for example ATCC VR-740 and those described in Hamre (1966) *Proc Soc Exp Biol Med* 121:190.

Delivery of the compositions of this invention into cells is not limited to the above mentioned viral vectors. Other delivery methods and media may be employed such as, for example, nucleic acid expression vectors, polycationic condensed DNA linked or unlinked to killed adenovirus alone, for example see US Serial No. 08/366,787, filed December 30, 1994 and Curiel (1992) *Hum Gene Ther* 3:147-154 ligand linked DNA, for example see Wu (1989) *J Biol Chem* 264:16985-16987, eucaryotic cell delivery vehicles cells, for example see US Serial No.08/240,030, filed May 9, 1994, and US Serial No. 08/404,796, deposition of photopolymerized hydrogel materials, hand-held gene transfer particle gun, as described in US Patent 5,149,655, ionizing radiation as described in US5,206,152 and in WO92/11033, nucleic charge neutralization or fusion with cell membranes. Additional approaches are described in Philip (1994) *Mol Cell Biol* 14:2411-2418 and in Woffendin (1994) *Proc Natl Acad Sci* 91:1581-1585.

Particle mediated gene transfer may be employed, for example see US Serial No. 60/023,867. Briefly, the sequence can be inserted into conventional vectors that contain conventional control sequences for high level expression, and then incubated with synthetic gene transfer molecules such as polymeric DNA-binding cations like polylysine, protamine, and albumin, linked to cell targeting ligands such as asialoorosomucoid, as described in Wu &

Wu (1987) *J. Biol. Chem.* 262:4429-4432, insulin as described in Hucked (1990) *Biochem Pharmacol* 40:253-263, galactose as described in Plank (1992) *Bioconjugate Chem* 3:533-539, lactose or transferrin.

Naked DNA may also be employed to transform a host cell. Exemplary naked DNA introduction methods are described in WO 90/11092 and US 5,580,859. Uptake efficiency may be improved using biodegradable latex beads. DNA coated latex beads are efficiently transported into cells after endocytosis initiation by the beads. The method may be improved further by treatment of the beads to increase hydrophobicity and thereby facilitate disruption of the endosome and release of the DNA into the cytoplasm.

Liposomes that can act as gene delivery vehicles are described in U.S. 5,422,120, WO95/13796, WO94/23697, WO91/14445 and EP-524,968. As described in USSN. 60/023,867, on non-viral delivery, the nucleic acid sequences encoding a polypeptide can be inserted into conventional vectors that contain conventional control sequences for high level expression, and then be incubated with synthetic gene transfer molecules such as polymeric DNA-binding cations like polylysine, protamine, and albumin, linked to cell targeting ligands such as asialoorosomucoid, insulin, galactose, lactose, or transferrin. Other delivery systems include the use of liposomes to encapsulate DNA comprising the gene under the control of a variety of tissue-specific or ubiquitously-active promoters. Further non-viral delivery suitable for use includes mechanical delivery systems such as the approach described in Woffendin *et al* (1994) *Proc. Natl. Acad. Sci. USA* 91(24):11581-11585. Moreover, the coding sequence and the product of expression of such can be delivered through deposition of photopolymerized hydrogel materials. Other conventional methods for gene delivery that can be used for delivery of the coding sequence include, for example, use of hand-held gene transfer particle gun, as described in U.S. 5,149,655; use of ionizing radiation for activating transferred gene, as described in U.S. 5,206,152 and WO92/11033.

Exemplary liposome and polycationic gene delivery vehicles are those described in US 5,422,120 and 4,762,915; inWO 95/13796; WO94/23697; and WO91/14445; in EP-0524968; and in Stryer, Biochemistry, pages 236-240 (1975) W.H. Freeman, San Francisco; Szoka (1980) *Biochem Biophys Acta* 600:1; Bayer (1979) *Biochem Biophys Acta* 550:464; Rivnay (1987) *Meth Enzymol* 149:119; Wang (1987) *Proc Natl Acad Sci* 84:7851; Plant (1989) *Anal Biochem* 176:420.

A polynucleotide composition can comprises therapeutically effective amount of a gene therapy vehicle, as the term is defined above. For purposes of the present invention, an effective dose will be from about 0.01 mg/ kg to 50 mg/kg or 0.05 mg/kg to about 10 mg/kg of the DNA constructs in the individual to which it is administered.

Delivery Methods

Once formulated, the polynucleotide compositions of the invention can be administered (1) directly to the subject; (2) delivered ex vivo, to cells derived from the subject; or (3) in vitro for expression of recombinant proteins. The subjects to be treated can be mammals or birds. Also, human subjects can be treated.

Direct delivery of the compositions will generally be accomplished by injection, either subcutaneously, intraperitoneally, intravenously or intramuscularly or delivered to the interstitial space of a tissue. The compositions can also be administered into a tumor or lesion. Other modes of administration include oral and pulmonary administration, suppositories, and transdermal applications, needles, and gene guns or hyposprays. Dosage treatment may be a single dose schedule or a multiple dose schedule.

Methods for the *ex vivo* delivery and reimplantation of transformed cells into a subject are known in the art and described in *eg.* WO93/14778. Examples of cells useful in ex vivo applications include, for example, stem cells, particularly hematopoetic, lymph cells, macrophages, dendritic cells, or tumor cells.

Generally, delivery of nucleic acids for both *ex vivo* and *in vitro* applications can be accomplished by the following procedures, for example, dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei, all well known in the art.

Polynucleotide and polypeptide pharmaceutical compositions

In addition to the pharmaceutically acceptable carriers and salts described above, the following additional agents can be used with polynucleotide and/or polypeptide compositions.

A.Polypeptides

One example are polypeptides which include, without limitation: asioloorosomucoid (ASOR); transferrin; asialoglycoproteins; antibodies; antibody fragments; ferritin; interleukins; interferons, granulocyte, macrophage colony stimulating factor (GM-CSF),

granulocyte colony stimulating factor (G-CSF), macrophage colony stimulating factor (M-CSF), stem cell factor and erythropoietin. Viral antigens, such as envelope proteins, can also be used. Also, proteins from other invasive organisms, such as the 17 amino acid peptide from the circumsporozoite protein of plasmodium falciparum known as RII.

B.Hormones, Vitamins, Etc.

Other groups that can be included are, for example: hormones, steroids, androgens, estrogens, thyroid hormone, or vitamins, folic acid.

C.Polyalkylenes, Polysaccharides, *etc.*

Also, polyalkylene glycol can be included with the desired polynucleotides or polypeptides. In a preferred embodiment, the polyalkylene glycol is polyethlylene glycol. In addition, mono-, di-, or polysaccarides can be included. In a preferred embodiment of this aspect, the polysaccharide is dextran or DEAE-dextran. Also, chitosan and poly(lactide-co-glycolide)

D.Lipids, and Liposomes

The desired polynucleotide or polypeptide can also be encapsulated in lipids or packaged in liposomes prior to delivery to the subject or to cells derived therefrom.

Lipid encapsulation is generally accomplished using liposomes which are able to stably bind or entrap and retain nucleic acid. The ratio of condensed polynucleotide or polypeptide to lipid preparation can vary but will generally be around 1:1 (mg DNA:micromoles lipid), or more of lipid. For a review of the use of liposomes as carriers for delivery of nucleic acids, see, Hug and Sleight (1991) *Biochim. Biophys. Acta.* 1097:1-17; Straubinger (1983) *Meth. Enzymol.* 101:512-527.

Liposomal preparations for use in the present invention include cationic (positively charged), anionic (negatively charged) and neutral preparations. Cationic liposomes have been shown to mediate intracellular delivery of plasmid DNA (Felgner (1987) *Proc. Natl. Acad. Sci. USA* 84:7413-7416); mRNA (Malone (1989) *Proc. Natl. Acad. Sci. USA* 86:6077-6081); and purified transcription factors (Debs (1990) *J. Biol. Chem.* 265:10189-10192), in functional form.

Cationic liposomes are readily available. For example, N[1-2,3-dioleyloxy)propyl]-N,N,N-triethylammonium (DOTMA) liposomes are available under the trademark Lipofectin, from GIBCO BRL, Grand Island, NY. (See, also, Felgner *supra*). Other commercially available liposomes include transfectace (DDAB/DOPE) and

DOTAP/DOPE (Boerhinger). Other cationic liposomes can be prepared from readily available materials using techniques well known in the art. See, *eg.* Szoka (1978) *Proc. Natl. Acad. Sci. USA* 75:4194-4198; WO90/11092 for a description of the synthesis of DOTAP (1,2-bis(oleoyloxy)-3-(trimethylammonio)propane) liposomes.

Similarly, anionic and neutral liposomes are readily available, such as from Avanti Polar Lipids (Birmingham, AL), or can be easily prepared using readily available materials. Such materials include phosphatidyl choline, cholesterol, phosphatidyl ethanolamine, dioleoylphosphatidyl choline (DOPC), dioleoylphosphatidyl glycerol (DOPG), dioleoylphoshatidyl ethanolamine (DOPE), among others. These materials can also be mixed with the DOTMA and DOTAP starting materials in appropriate ratios. Methods for making liposomes using these materials are well known in the art.

The liposomes can comprise multilammelar vesicles (MLVs), small unilamellar vesicles (SUVs), or large unilamellar vesicles (LUVs). The various liposome-nucleic acid complexes are prepared using methods known in the art. See *eg.* Straubinger (1983) *Meth. Immunol.* 101:512-527; Szoka (1978) *Proc. Natl. Acad. Sci. USA* 75:4194-4198; Papahadjopoulos (1975) *Biochim. Biophys. Acta* 394:483; Wilson (1979) *Cell* 17:77); Deamer & Bangham (1976) *Biochim. Biophys. Acta* 443:629; Ostro (1977) *Biochem. Biophys. Res. Commun.* 76:836; Fraley (1979) *Proc. Natl. Acad. Sci. USA* 76:3348); Enoch & Strittmatter (1979) *Proc. Natl. Acad. Sci. USA* 76:145; Fraley (1980) *J. Biol. Chem.* (1980) 255:10431; Szoka & Papahadjopoulos (1978) *Proc. Natl. Acad. Sci. USA* 75:145; and Schaefer-Ridder (1982) *Science* 215:166.

E.Lipoproteins

In addition, lipoproteins can be included with the polynucleotide or polypeptide to be delivered. Examples of lipoproteins to be utilized include: chylomicrons, HDL, IDL, LDL, and VLDL. Mutants, fragments, or fusions of these proteins can also be used. Also, modifications of naturally occurring lipoproteins can be used, such as acetylated LDL. These lipoproteins can target the delivery of polynucleotides to cells expressing lipoprotein receptors. Preferably, if lipoproteins are including with the polynucleotide to be delivered, no other targeting ligand is included in the composition.

Naturally occurring lipoproteins comprise a lipid and a protein portion. The protein portion are known as apoproteins. At the present, apoproteins A, B, C, D, and E have been

isolated and identified. At least two of these contain several proteins, designated by Roman numerals, AI, AII, AIV; CI, CII, CIII.

A lipoprotein can comprise more than one apoprotein. For example, naturally occurring chylomicrons comprises of A, B, C, and E, over time these lipoproteins lose A and acquire C and E apoproteins. VLDL comprises A, B, C, and E apoproteins, LDL comprises apoprotein B; and HDL comprises apoproteins A, C, and E.

The amino acid of these apoproteins are known and are described in, for example, Breslow (1985) *Annu Rev. Biochem* 54:699; Law (1986) *Adv. Exp Med. Biol.* 151:162; Chen (1986) *J Biol Chem* 261:12918; Kane (1980) *Proc Natl Acad Sci USA* 77:2465; and Utermann (1984) *Hum Genet* 65:232.

Lipoproteins contain a variety of lipids including, triglycerides, cholesterol (free and esters), and phopholipids. The composition of the lipids varies in naturally occurring lipoproteins. For example, chylomicrons comprise mainly triglycerides. A more detailed description of the lipid content of naturally occurring lipoproteins can be found, for example, in *Meth. Enzymol.* 128 (1986). The composition of the lipids are chosen to aid in conformation of the apoprotein for receptor binding activity. The composition of lipids can also be chosen to facilitate hydrophobic interaction and association with the polynucleotide binding molecule.

Naturally occurring lipoproteins can be isolated from serum by ultracentrifugation, for instance. Such methods are described in *Meth. Enzymol. (supra)*; Pitas (1980) *J. Biochem.* 255:5454-5460 and Mahey (1979) *J Clin. Invest* 64:743-750.

Lipoproteins can also be produced by *in vitro* or recombinant methods by expression of the apoprotein genes in a desired host cell. See, for example, Atkinson (1986) *Annu Rev Biophys Chem* 15:403 and Radding (1958) *Biochim Biophys Acta* 30: 443.

Lipoproteins can also be purchased from commercial suppliers, such as Biomedical Techniologies, Inc., Stoughton, Massachusetts, USA.

Further description of lipoproteins can be found in Zuckermann et al., PCT. Appln. No. US97/14465.

F.Polycationic Agents

Polycationic agents can be included, with or without lipoprotein, in a composition with the desired polynucleotide or polypeptide to be delivered.

Polycationic agents, typically, exhibit a net positive charge at physiological relevant pH and are capable of neutralizing the electrical charge of nucleic acids to facilitate delivery to a desired location. These agents have both in vitro, ex vivo, and in vivo applications. Polycationic agents can be used to deliver nucleic acids to a living subject either intramuscularly, subcutaneously, etc.

The following are examples of useful polypeptides as polycationic agents: polylysine, polyarginine, polyornithine, and protamine. Other examples include histones, protamines, human serum albumin, DNA binding proteins, non-histone chromosomal proteins, coat proteins from DNA viruses, such as (X174, transcriptional factors also contain domains that bind DNA and therefore may be useful as nucleic aid condensing agents. Briefly, transcriptional factors such as C/CEBP, c-jun, c-fos, AP-1, AP-2, AP-3, CPF, Prot-1, Sp-1, Oct-1, Oct-2, CREP, and TFIID contain basic domains that bind DNA sequences.

Organic polycationic agents include: spermine, spermidine, and purtrescine.

The dimensions and of the physical properties of a polycationic agent can be extrapolated from the list above, to construct other polypeptide polycationic agents or to produce synthetic polycationic agents.

Synthetic Polycationic Agents

Synthetic polycationic agents which are useful include, for example, DEAE-dextran, polybrene. Lipofectin□, and lipofectAMINE□ are monomers that form polycationic complexes when combined with polynucleotides or polypeptides.

Immunodiagnostic Assays

Neisserial antigens of the invention can be used in immunoassays to detect antibody levels (or, conversely, anti-Neisserial antibodies can be used to detect antigen levels). Immunoassays based on well defined, recombinant antigens can be developed to replace invasive diagnostics methods. Antibodies to Neisserial proteins within biological samples, including for example, blood or serum samples, can be detected. Design of the immunoassays is subject to a great deal of variation, and a variety of these are known in the art. Protocols for the immunoassay may be based, for example, upon competition, or direct reaction, or sandwich type assays. Protocols may also, for example, use solid supports, or may be by immunoprecipitation. Most assays involve the use of labeled antibody or polypeptide; the labels may be, for example, fluorescent, chemiluminescent, radioactive, or dye molecules. Assays which amplify the signals from the probe are also known; examples of which are

assays which utilize biotin and avidin, and enzyme-labeled and mediated immunoassays, such as ELISA assays.

Kits suitable for immunodiagnosis and containing the appropriate labeled reagents are constructed by packaging the appropriate materials, including the compositions of the invention, in suitable containers, along with the remaining reagents and materials (for example, suitable buffers, salt solutions, *etc.*) required for the conduct of the assay, as well as suitable set of assay instructions.

Nucleic Acid Hybridisation

"Hybridization" refers to the association of two nucleic acid sequences to one another by hydrogen bonding. Typically, one sequence will be fixed to a solid support and the other will be free in solution. Then, the two sequences will be placed in contact with one another under conditions that favor hydrogen bonding. Factors that affect this bonding include: the type and volume of solvent; reaction temperature; time of hybridization; agitation; agents to block the non-specific attachment of the liquid phase sequence to the solid support (Denhardt's reagent or BLOTTO); concentration of the sequences; use of compounds to increase the rate of association of sequences (dextran sulfate or polyethylene glycol); and the stringency of the washing conditions following hybridization. See Sambrook *et al.* [*supra*] Volume 2, chapter 9, pages 9.47 to 9.57.

"Stringency" refers to conditions in a hybridization reaction that favor association of very similar sequences over sequences that differ. For example, the combination of temperature and salt concentration should be chosen that is approximately 120 to 200□C below the calculated Tm of the hybrid under study. The temperature and salt conditions can often be determined empirically in preliminary experiments in which samples of genomic DNA immobilized on filters are hybridized to the sequence of interest and then washed under conditions of different stringencies. See Sambrook *et al.* at page 9.50.

Variables to consider when performing, for example, a Southern blot are (1) the complexity of the DNA being blotted and (2) the homology between the probe and the sequences being detected. The total amount of the fragment(s) to be studied can vary a magnitude of 10, from 0.1 to 1μg for a plasmid or phage digest to 10^{-9} to 10^{-8} g for a single copy gene in a highly complex eukaryotic genome. For lower complexity polynucleotides, substantially shorter blotting, hybridization, and exposure times, a smaller amount of starting polynucleotides, and lower specific activity of probes can be used. For example, a single-copy

yeast gene can be detected with an exposure time of only 1 hour starting with 1 μg of yeast DNA, blotting for two hours, and hybridizing for 4-8 hours with a probe of 10^8 cpm/μg. For a single-copy mammalian gene a conservative approach would start with 10 μg of DNA, blot overnight, and hybridize overnight in the presence of 10% dextran sulfate using a probe of greater than 10^8 cpm/μg, resulting in an exposure time of ~24 hours.

Several factors can affect the melting temperature (Tm) of a DNA-DNA hybrid between the probe and the fragment of interest, and consequently, the appropriate conditions for hybridization and washing. In many cases the probe is not 100% homologous to the fragment. Other commonly encountered variables include the length and total G+C content of the hybridizing sequences and the ionic strength and formamide content of the hybridization buffer. The effects of all of these factors can be approximated by a single equation:

$$Tm= 81 + 16.6(\log_{10}Ci) + 0.4[\%(G + C)]-0.6(\%formamide) - 600/n-1.5(\%mismatch).$$

where Ci is the salt concentration (monovalent ions) and n is the length of the hybrid in base pairs (slightly modified from Meinkoth & Wahl (1984) *Anal. Biochem.* 138: 267-284).

In designing a hybridization experiment, some factors affecting nucleic acid hybridization can be conveniently altered. The temperature of the hybridization and washes and the salt concentration during the washes are the simplest to adjust. As the temperature of the hybridization increases (*ie.* stringency), it becomes less likely for hybridization to occur between strands that are nonhomologous, and as a result, background decreases. If the radiolabeled probe is not completely homologous with the immobilized fragment (as is frequently the case in gene family and interspecies hybridization experiments), the hybridization temperature must be reduced, and background will increase. The temperature of the washes affects the intensity of the hybridizing band and the degree of background in a similar manner. The stringency of the washes is also increased with decreasing salt concentrations.

In general, convenient hybridization temperatures in the presence of 50% formamide are 42□C for a probe with is 95% to 100% homologous to the target fragment, 37□C for 90% to 95% homology, and 32□C for 85% to 90% homology. For lower homologies, formamide content should be lowered and temperature adjusted accordingly, using the equation above. If the homology between the probe and the target fragment are not known, the simplest approach is to start with both hybridization and wash conditions which are nonstringent. If non-specific bands or high background are observed after autoradiography, the filter can be

washed at high stringency and reexposed. If the time required for exposure makes this approach impractical, several hybridization and/or washing stringencies should be tested in parallel.

Nucleic Acid Probe Assays

Methods such as PCR, branched DNA probe assays, or blotting techniques utilizing nucleic acid probes according to the invention can determine the presence of cDNA or mRNA. A probe is said to "hybridize" with a sequence of the invention if it can form a duplex or double stranded complex, which is stable enough to be detected.

The nucleic acid probes will hybridize to the Neisserial nucleotide sequences of the invention (including both sense and antisense strands). Though many different nucleotide sequences will encode the amino acid sequence, the native Neisserial sequence is preferred because it is the actual sequence present in cells. mRNA represents a coding sequence and so a probe should be complementary to the coding sequence; single-stranded cDNA is complementary to mRNA, and so a cDNA probe should be complementary to the non-coding sequence.

The probe sequence need not be identical to the Neisserial sequence (or its complement) — some variation in the sequence and length can lead to increased assay sensitivity if the nucleic acid probe can form a duplex with target nucleotides, which can be detected. Also, the nucleic acid probe can include additional nucleotides to stabilize the formed duplex. Additional Neisserial sequence may also be helpful as a label to detect the formed duplex. For example, a non-complementary nucleotide sequence may be attached to the 5' end of the probe, with the remainder of the probe sequence being complementary to a Neisserial sequence. Alternatively, non-complementary bases or longer sequences can be interspersed into the probe, provided that the probe sequence has sufficient complementarity with the a Neisserial sequence in order to hybridize therewith and thereby form a duplex which can be detected.

The exact length and sequence of the probe will depend on the hybridization conditions, such as temperature, salt condition and the like. For example, for diagnostic applications, depending on the complexity of the analyte sequence, the nucleic acid probe typically contains at least 10-20 nucleotides, preferably 15-25, and more preferably at least 30 nucleotides, although it may be shorter than this. Short primers generally require cooler temperatures to form sufficiently stable hybrid complexes with the template.

Probes may be produced by synthetic procedures, such as the triester method of Matteucci *et al.* [*J. Am. Chem. Soc.* (1981) 103:3185], or according to Urdea *et al.* [*Proc. Natl. Acad. Sci. USA* (1983) 80: 7461], or using commercially available automated oligonucleotide synthesizers.

The chemical nature of the probe can be selected according to preference. For certain applications, DNA or RNA are appropriate. For other applications, modifications may be incorporated *eg.* backbone modifications, such as phosphorothioates or methylphosphonates, can be used to increase *in vivo* half-life, alter RNA affinity, increase nuclease resistance *etc.* [*eg.* see Agrawal & Iyer (1995) *Curr Opin Biotechnol* 6:12-19; Agrawal (1996) *TIBTECH* 14:376-387]; analogues such as peptide nucleic acids may also be used [*eg.* see Corey (1997) *TIBTECH* 15:224-229; Buchardt *et al.* (1993) *TIBTECH* 11:384-386].

One example of a nucleotide hybridization assay is described by Urdea *et al.* in international patent application WO92/02526 [see also US patent 5,124,246].

Alternatively, the polymerase chain reaction (PCR) is another well-known means for detecting small amounts of target nucleic acids. The assay is described in: Mullis *et al.* [*Meth. Enzymol.* (1987) 155: 335-350]; US patent 4,683,195; and US patent 4,683,202. Two "primer" nucleotides hybridize with the target nucleic acids and are used to prime the reaction. The primers can comprise sequence that does not hybridize to the sequence of the amplification target (or its complement) to aid with duplex stability or, for example, to incorporate a convenient restriction site. Typically, such sequence will flank the desired Neisserial sequence.

A thermostable polymerase creates copies of target nucleic acids from the primers using the original target nucleic acids as a template. After a threshold amount of target nucleic acids are generated by the polymerase, they can be detected by more traditional methods, such as Southern blots. When using the Southern blot method, the labelled probe will hybridize to the Neisserial sequence (or its complement).

Also, mRNA or cDNA can be detected by traditional blotting techniques described in Sambrook *et al* [*supra*]. mRNA, or cDNA generated from mRNA using a polymerase enzyme, can be purified and separated using gel electrophoresis. The nucleic acids on the gel are then blotted onto a solid support, such as nitrocellulose. The solid support is exposed to a labelled probe and then washed to remove any unhybridized probe. Next, the duplexes

containing the labeled probe are detected. Typically, the probe is labelled with a radioactive moiety.

EXAMPLES

The examples describe nucleic acid sequences which have been identified in *N. meningitidis*, and *N. gonorrhoeae* along with their respective and putative translation products. Not all of the nucleic acid sequences are complete *ie.* they encode less than the full-length wild-type protein.

The examples are generally in the following format:

- a nucleotide sequence which has been identified in *N. meningitidis*
- the putative translation product of said *N. meningitidis* sequence
- a computer analysis of said translation product based on database comparisons
- a corresponding nucleotide sequence identified from *N. gonorrhoeae*
- the putative translation product of said *N. gonorrhoeae* sequence
- a comparision of the percentage of identity between the translation product of the *N. meningitidis* sequence and the *N. gonorrhoeae* sequence.
- a corresponding nucleotide sequence identified from strain A of *N. meningitidis*
- the putative translation product of said *N. meningitidis* strain A sequence
- a comparision of the percentage of identity between the translation product of the *N. meningitidis* sequence and the *N. gonorrhoeae* sequence.
- a description of the characteristics of the protein which indicates that it might be suitably antigenic or immunogenic.

Sequence comparisons were performed at NCBI (http://www.ncbi.nlm.nih.gov) using the algorithms BLAST, BLAST2, BLASTn, BLASTp, tBLASTn, BLASTx, & tBLASTx [*eg.* see also Altschul *et al.* (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25:2289-3402]. Searches were performed against the following databases: non-redundant GenBank+EMBL+DDBJ+PDB sequences and non-redundant GenBank CDS translations+PDB+SwissProt+SPupdate+PIR sequences.

Dots within nucleotide sequences represent nucleotides which have been arbitrarily introduced in order to maintain a reading frame. In the same way, double-underlined nucleotides were removed . Lower case letters represent ambiguities which arose during

alignment of independent sequencing reactions (some of the nucleotide sequences in the examples are derived from combining the results of two or more experiments).

Nucleotide sequences were scanned in all six reading frames to predict the presence of hydrophobic domains using an algorithm based on the statistical studies of Esposti *et al.* [Critical evaluation of the hydropathy of membrane proteins (1990) *Eur J Biochem* 190:207-219]. These domains represent potential transmembrane regions or hydrophobic leader sequences.

Open reading frames were predicted from fragmented nucleotide sequences using the program ORFFINDER (NCBI).

Underlined amino acid sequences indicate possible transmembrane domains or leader sequences in the ORFs, as predicted by the PSORT algorithm (http://www.psort.nibb.ac.jp). Functional domains were also predicted using the MOTIFS program (GCG Wisconsin & PROSITE).

For each of the following examples: based on the presence of a putative leader sequence and/or several putative transmembrane domains (single-underlined) in the gonococcal protein, it is predicted that the proteins from *N. meningitidis* and *N. gonorrhoeae*, and their respective epitopes, could be useful antigens or immunogenic compositions for vaccines or diagnostics.

The standard techniques and procedures which may be employed in order to perform the invention (*e.g.* to utilize the disclosed sequences for vaccination or diagnostic purposes) were summarized above. This summary is not a limitation on the invention but, rather, gives examples that may be used, but are not required.

In particular, the following methods were used to express, purify and biochemically characterize the proteins of the invention.

Chromosomal DNA Preparation

N.meningitidis strain 2996 was grown to exponential phase in 100ml of GC medium, harvested by centrifugation, and resuspended in 5ml buffer (20%(w/v) Sucrose, 50mM Tris-HCl, 50mM EDTA, pH8). After 10 minutes incubation on ice, the bacteria were lysed by adding 10ml of lysis solution (50mM NaCl, 1% Na-Sarkosyl, 50μg/ml Proteinase K), and the suspension incubated at 37°C for 2 hours. Two phenol extractions (equilibrated to pH 8) and one $CHCl_3$/isoamylalcohol (24:1) extraction were performed. DNA was precipitated by addition of 0.3M sodium acetate and 2 volumes of ethanol, and collected by centrifugation.

The pellet was washed once with 70%(v/v) ethanol and redissolved in 4.0ml TE buffer (10mM Tris-HCl, 1mM EDTA, pH 8.0). The DNA concentration was measured by reading the OD at 260 nm.

Oligonucleotide design

Synthetic oligonucleotide primers were designed on the basis of the coding sequence of each ORF, using (a) the meningococcus B sequence when available, or (b) the gonococcus/meningococcus A sequence, adapted to the codon preference usage of meningococcus as necessary. Any predicted signal peptides were omitted, by designing the 5' primers to sequence immediately downstream from the predicted leader sequence.

For most ORFs, the 5' primers included two restriction enzyme recognition sites (*BamH*I-*Nde*I, *BamH*I-*Nhe*I, *EcoR*I-*Nde*I or *EcoR*I-*Nhe*I), depending on the restriction pattern of the gene of interest. The 3' primers included a *Xho*I or a *Hind*III restriction site (table 1). This procedure was established in order to direct the cloning of each amplification product (corresponding to each ORF) into two different expression systems: pGEX-KG (using *BamH*I-*Xho*I, *BamH*I-*Hind*III, *EcoR*I-*Xho*I or *EcoR*I-*Hind*III), and pET21b+ (using *Nde*I-*Xho*I, *Nhe*I-*Xho*I, *Nde*I-*Hind*III or *Nhe*I-*Hind*III).

5'-end primer tail:	CGCGGATCCCATATG	(*BamH*I-*Nde*I)
	CGCGGATCCGCTAGC	(*BamH*I-*Nhe*I)
	CCGGAATTCTACATATG	(*EcoR*I-*Nde*I)
	CCGGAATTCTAGCTAGC	(*EcoR*I-*Nhe*I)
3'-end primer tail:	CCCGCTCGAG	(*Xho*I)
	CCCGCTCGAG	(*Hind*III)

For cloning ORFs into the pGEX-His vector, the 5' and 3' primers contained only one restriction enzyme site (*EcoR*I, *Kpn*I or *Sal*I for the 5' primers and *Pst*I, *Xba*I, *Sph*I or *Sal*I for the 3' primers). Again restriction sites were chosen according to the particular restriction pattern of the gene (table 1).

5'-end primer tail:	(AAA) AAAGAATTC	(*EcoR*I)
	(AAA) AAAGGTACC	(*Kpn*I)
3'-end primer tail:	(AAA) AAACTGCAG	(*Pst*I)
	(AAA) AAATCTAGA	(*Xba*I)

AAAGCATGC (*Sph*I)

5' or 3'-end primer tail: AAAAAAGTCGAC (*Sal*I)

As well as containing the restriction enzyme recognition sequences, the primers included nucleotides which hybridized to the sequence to be amplified. The melting temperature depended on the number and type of hybridising nucleotides in the whole primer, and was determined for each primer using the formulae:

$T_m = 4\,(G+C) + 2\,(A+T)$ (tail excluded)

$T_m = 64.9 + 0.41\,(\%\,GC) - 600/N$ (whole primer)

The melting temperatures of the selected oligonucleotides were usually 65-70°C for the whole oligo and 50-55°C for the hybridising region alone.

Table 1 shows the forward and reverse primers used for each amplification. In certain cases, the sequence of the primer does not exactly match the sequence of the predicted ORF. This is because when initial amplifications were performed, the complete 5' and/or 3' sequences for some meningococcal B ORFs were not known. However the corresponding sequences had been identified in Gonococcus or in Meningoccus A. Hence, when the Meningoccus B sequence was incomplete or uncertain, Gonococcal or Meningococcal A sequences were used as the basis for primer design. These sequences were altered to take account of codon preference. It can be appreciated that, once the complete sequence is identified, this approach will no longer be necessary.

Oligonucleotides were synthesized using a Perkin Elmer 394 DNA/RNA Synthesizer, eluted from the columns in 2.0ml NH_4OH, and deprotected by 5 hours incubation at 56°C. The oligos were precipitated by addition of 0.3M Na-Acetate and 2 volumes ethanol. The samples were centrifuged and the pellets resuspended in either 100μl or 1.0ml of water. The OD_{260} was determined using a Perkin Elmer Lambda Bio spectophotometer and the concentration adjusted to 2-10pmol/μl.

Amplification

The standard PCR protocol was as follows: 50-200ng of genomic DNA was used as a template in the presence of 20-40μM of each oligonucletide primer, 400-800μM dNTPs solution, 1x PCR buffer (including 1.5mM $MgCl_2$), 2.5 units *TaqI* DNA polymerase (using

Perkin-Elmer AmpliTaQ, GIBCO Platinum, Pwo DNA polymerase, or Tahara Shuzo Taq polymerase). In some cases, PCR was optimsed by the addition of 10μl DMSO or 50μl 2M Betaine.

After a hot start (adding the polymerase during a preliminary 3 minute incubation of the whole mix at 95°C), each sample underwent a two-step amplification. The first 5 cycles were performed using the hybridization temperature that excluded the restriction enzyme tail of the primer (see above). This was followed by 30 cycles using the hybridization temperature calculated for the whole length oligos. The cycles were completed with a 10 minute extension step at 72°C. The standard cycles were as follows:

	Denaturation	**Hybridisation**	**Elongation**
First 5 cycles	30 seconds 95°C	30 seconds 50-55°C	30-60 seconds 72°C
Last 30 cycles	30 seconds 95°C	30 seconds 65-70°C	30-60 seconds 72°C

Elongation times varied according to the length of the ORF to be amplified. Amplifications were performed using either a 9600 or a 2400 Perkin Elmer GeneAmp PCR System. To check the results, 1/10 of the amplification volume was loaded onto a 1-1.5% (w/v) agarose gel and the size of each amplified fragment compared with a DNA molecular weight marker.

The amplified DNA was either loaded directly on a 1% agarose gel or first precipitated with ethanol and resuspended in a volume suitable to be loaded on a 1.0% agarose gel. The DNA fragment corresponding to the band of correct size was purified using the Qiagen Gel Extraction Kit, following the manufacturer's protocol. DNA fragments were eluted in a volume of 30μl or 50μl with either H2O or 10mM Tris, pH 8.5.

Digestion of PCR fragments

The purified DNA corresponding to the amplified fragment was doubly-digested with the appropriate restriction enzymes for; cloning into pET-21b+ and expressing the protein as a C-terminus His-tagged fusion, for cloning into pGEX-KG and expressing the protein as a N-

terminus GST-fusion, and for cloning into pGEX-His and expressing the protein as a N-terminus GST-His tagged fusion.

Each purified DNA fragment was incubated at 37°C for 3 hours to overnight with 20 units of appropriate restriction enzyme (New England Biolabs) in a volume of either 30 or 40μl in the presence of suitable digestion buffer. Digested fragments were purified using the QIAquick PCR purification kit (following the manufacturer's instructions) and eluted in a volume of 30μl or 50μl with either H2O or 10mM Tris, pH 8.5. The DNA concentration was determined by quantitative agarose gel electrophoresis (1.0% gel) in the presence of a titrated molecular weight marker.

Digestion of the cloning vectors (pET22B, pGEX-KG, pTRC-His A, pET21b+, pGEX-KG, and pGEX-His)

The vector pGEX-His is a modified pGEX-2T vector carrying a region encoding six histidine residues upstream of the thrombin cleavage site and containing the multiple cloning site of the vector pTRC99 (Pharmacia).10 μg plasmid was double-digested with 50 units of each restriction enzyme in 200 μl reaction volume in the presence of appropriate buffer by overnight incubation at 37°C. After loading the whole digestion on a 1% agarose gel, the band corresponding to the digested vector was purified from the gel using the Qiagen QIAquick Gel Extraction Kit and the DNA was eluted in 50 μl of 10 mM Tris-HCl, pH 8.5. The DNA concentration was evaluated by measuring OD_{260} of the sample, and adjusted to 50 μg/μl. 1 μl of plasmid was used for each cloning procedure.

10μg of plasmid vector was doubly-digested with 50 units of each restriction enzyme in a volume of 200μl with the appropriate buffer overnight at 37°C. The digest was loaded onto a 1.0% agarose gel and the band corresponding to the digested vector purified using the Qiagen QIAquick Gel Extraction Kit. DNA was eluted in 50μl of 10mM Tris-HCl, pH 8.5. The DNA concentration was evaluated by measuring OD_{260nm} and the concentration adjusted to 50μg/μl. 1μl of plasmid was used for each cloning procedure.

Cloning

For some ORFs, the fragments corresponding to each ORF, previously digested and purified, were ligated in both pET22b and pGEX-KG. In a final volume of 20 μl, a molar

ratio of 3:1 fragment/vector was ligated using 0.5 µl of NEB T4 DNA ligase (400 units/µl), in the presence of the buffer supplied by the manufacturer. The reaction was incubated at room temperature for 3 hours. In some experiments, ligation was performed using the Boheringer "Rapid Ligation Kit", following the manufacturer's instructions.

In order to introduce the recombinant plasmid in a suitable strain, 100 µl *E. coli* DH5 competent cells were incubated with the ligase reaction solution for 40 minutes on ice, then at 37°C for 3 minutes, then, after adding 800 µl LB broth, again at 37°C for 20 minutes. The cells were then centrifuged at maximum speed in an Eppendorf microfuge and resuspended in approximately 200 µl of the supernatant. The suspension was then plated on LB ampicillin (100 mg/ml).

The screening of the recombinant clones was performed by growing 5 randomly-chosen colonies overnight at 37 °C in either 2 ml (pGEX or pTC clones) or 5ml (pET clones) LB broth + 100 µg/ml ampicillin. The cells were then pelletted and the DNA extracted using the Qiagen QIAprep Spin Miniprep Kit, following the manufacturer's instructions, to a final volume of 30 µl. 5 µl of each individual miniprep (approximately 1g) were digested with either *Nde*I/*Xho*I or *BamH*I/*Xho*I and the whole digestion loaded onto a 1-1.5% agarose gel (depending on the expected insert size), in parallel with the molecular weight marker (1Kb DNA Ladder, GIBCO). The screening of the positive clones was made on the base of the correct insert size.

For other ORFs, the fragments corresponding to each ORF, previously digested and purified, were ligated into both pET21b+ and pGEX-KG. A molar ratio of of 3:1 fragment/vector was used in a final volume of 20µl, that included 0.5µl T4 DNA ligase (400 units/µl, NEB) and ligation buffer supplied by the manufacturer. The reaction was performed at room temperature for 3 hours. In some experiments, ligation was performed using the Boheringer "Rapid Ligation Kit" and the manufacturer's protocol.

Recombinant plasmid was transformed into 100µl of competent *E. coli* DH5 or HB101 by incubating the ligase reaction solution and bacteria for 40 minutes on ice then at 37°C for 3 minutes. This was followed by the addition of 800µl LB broth and incubation at 37°C for 20 minutes. The cells were centrifuged at maximum speed in an Eppendorf microfuge, resuspended in approximately 200µl of the supernatant and plated onto LB ampicillin (100mg/ml) agar.

Screening for recombinant clones was performed by growing 5 randomly selected colonies overnight at 37°C in either 2.0ml (pGEX-KG clones) or 5.0ml (pET clones) LB broth + 100μg/ml ampicillin. Cells were pelleted and plasmid DNA extracted using the Qiagen QIAprep Spin Miniprep Kit, following the manufacturer's instructions. Approximately 1μg of each individual miniprep was digested with the appropriate restriction enzymes and the digest loaded onto a 1-1.5% agarose gel (depending on the expected insert size), in parallel with the molecular weight marker (1kb DNA Ladder, GIBCO). Positive clones were selected on the basis of the size of insert.

ORFs were cloned into PGEX-His, by doubly-digesting the PCR product and ligating into similarly digested vector. After cloning, recombinant plasmids were transformed into the *E.coli* host W3110. Individual clones were grown overnight at 37°C in LB broth with 50μg/ml ampicillin.

Certain ORFs may be cloned into the pGEX-HIS vector using *EcoR*I-*Pst*I cloning sites, or *EcoR*I-*Sal*I, or *Sal*I-*Pst*I. After cloning, the recombinant plasmids may be introduced in the *E.*coli host W3110.

Expression

Each ORF cloned into the expression vector may then be transformed into the strain suitable for expression of the recombinant protein product. 1 μl of each construct was used to transform 30 μl of *E.coli* BL21 (pGEX vector), *E.coli* TOP 10 (pTRC vector) or *E.coli* BL21-DE3 (pET vector), as described above. In the case of the pGEX-His vector, the same *E.coli* strain (W3110) was used for initial cloning and expression. Single recombinant colonies were inoculated into 2ml LB+Amp (100 μg/ml), incubated at 37°C overnight, then diluted 1:30 in 20 ml of LB+Amp (100 μg/ml) in 100 ml flasks, making sure that the OD_{600} ranged between 0.1 and 0.15. The flasks were incubated at 30°C into gyratory water bath shakers until OD indicated exponential growth suitable for induction of expression (0.4-0.8 OD for pET and pTRC vectors; 0.8-1 OD for pGEX and pGEX-His vectors). For the pET, pTRC and pGEX-His vectors, the protein expression was induced by addiction of 1mM IPTG, whereas in the case of pGEX system the final concentration of IPTG was 0.2 mM. After 3 hours incubation at 30°C, the final concentration of the sample was checked by OD. In order to check expression, 1ml of each sample was removed, centrifuged in a microfuge, the pellet

resuspended in PBS, and analysed by 12% SDS-PAGE with Coomassie Blue staining. The whole sample was centrifuged at 6000*g* and the pellet resuspended in PBS for further use.

GST-fusion proteins large-scale purification.

For some ORFs, a single colony was grown overnight at 37°C on LB+Amp agar plate. The bacteria were inoculated into 20 ml of LB+Amp liquid colture in a water bath shaker and grown overnight. Bacteria were diluted 1:30 into 600 ml of fresh medium and allowed to grow at the optimal temperature (20-37°C) to OD_{550} 0.8-1. Protein expression was induced with 0.2mM IPTG followed by three hours incubation. The culture was centrifuged at 8000 rpm at 4°C. The supernatant was discarded and the bacterial pellet was resuspended in 7.5 ml cold PBS. The cells were disrupted by sonication on ice for 30 sec at 40W using a Branson sonifier B-15, frozen and thawed two times and centrifuged again. The supernatant was collected and mixed with 150μl Glutatione-Sepharose 4B resin (Pharmacia) (previously washed with PBS) and incubated at room temperature for 30 minutes. The sample was centrifuged at 700*g* for 5 minutes at 4C. The resin was washed twice with 10 ml cold PBS for 10 minutes, resuspended in 1ml cold PBS, and loaded on a disposable column. The resin was washed twice with 2ml cold PBS until the flow-through reached OD_{280} of 0.02-0.06. The GST-fusion protein was eluted by addition of 700μl cold Glutathione elution buffer 10mM reduced glutathione, 50mM Tris-HCl) and fractions collected until the OD_{280} was 0.1. 21μl of each fraction were loaded on a 12% SDS gel using either Biorad SDS-PAGE Molecular weight standard broad range (M1) (200, 116.25, 97.4, 66.2, 45, 31, 21.5, 14.4, 6.5 kDa) or Amersham Rainbow Marker (M") (220, 66, 46, 30, 21.5, 14.3 kDa) as standards. As the MW of GST is 26kDa, this value must be added to the MW of each GST-fusion protein.

For other ORFs, for each clone to be purified as a GST-fusion, a single colony was streaked out and grown overnight at 37°C on a LB/Amp. (100μg/ml) agar plate. An isolated colony from this plate was inoculated into 20ml of LB/Amp (100 μg/ml) liquid medium and grown overnight at 37°C with shaking. The overnight culture was diluted 1:30 into 600ml LB/Amp (100μg/ml) liquid medium and allowed to grow at the optimal temperature (20-37°C) until the OD_{550nm} reached 0.6-0.8. Recombinant protein expression was induced by addition of IPTG (final concentration 0.2mM) and the culture incubated for a further 3 hours. Bacteria were harvested by centrifugation at 8000xg for 15 min at 4°C.

The bacterial pellet was resuspended in 7.5ml cold PBS. Cells were disrupted by sonication on ice four times for 30 sec at 40W using a Branson sonifier 450 and centrifuged at 13 000xg for 30 min at 4°C. The supernatant was collected and mixed with 150μl Glutatione-Sepharose 4B resin (Pharmacia), previously equilibrated with PBS, and incubated at room temperature with gentle agitation for 30 min. The batch-wise preparation was centrifuged at 700xg for 5 min at 4°C and the supernatant discarded. The resin was washed twice (batchwise) with 10ml cold PBS for 10 min, resuspended in 1ml cold PBS, and loaded onto a disposable column. The resin continued to be washed with cold PBS, until the OD_{280nm} of the flow-through reached 0.02-0.01. The GST-fusion protein was eluted by addition of 700μl cold glutathione elution buffer (10mM reduced glutathione, 50mM Tris-HCl pH 8.0) and fractions collected, until the OD_{280nm} of the eluate indicated all the recombinant protein was obtained. 20μl aliquots of each elution fraction were analyzed by SDS-PAGE using a 12% gel. The molecular mass of the purified proteins was determined using either the Bio-Rad broad range molecular weight standard (M1) (200, 116, 97.4, 66.2, 45.0, 31.0, 21.5, 14.4, 6.5 kDa) or the Amersham Rainbow Marker (M2) (220, 66.2, 46.0, 30.0, 21.5, 14.3 kDa). The molecular weights of GST-fusion proteins are a combination of the 26 kDa GST protein and its fusion partner. Protein concentrations were estimated using the Bradford assay.

His-fusion soluble proteins large-scale purification.

For some ORFs, a single colony was grown overnight at 37°C on a LB + Amp agar plate. The bacteria were inoculated into 20ml of LB+Amp liquid culture and incubated overnight in a water bath shaker. Bacteria were diluted 1:30 into 600ml fresh medium and allowed to grow at the optimal temperature (20-37°C) to OD_{550} 0.6-0.8. Protein expression was induced by addition of 1 mM IPTG and the culture further incubated for three hours. The culture was centrifuged at 8000 rpm at 4°C, the supernatant was discarded and the bacterial pellet was resuspended in 7.5ml cold 10mM imidazole buffer (300 mM NaCl, 50 mM phosphate buffer, 10 mM imidazole, pH 8). The cells were disrupted by sonication on ice for 30 sec at 40W using a Branson sonifier B-15, frozen and thawed two times and centrifuged again. The supernatant was collected and mixed with 150μl Ni^{2+}-resin (Pharmacia) (previously washed with 10mM imidazole buffer) and incubated at room temperature with gentle agitation for 30 minutes. The sample was centrifuged at 700g for 5 minutes at 4°C. The resin was washed twice with 10 ml cold 10mM imidazole buffer for 10 minutes,

resuspended in 1ml cold 10mM imidazole buffer and loaded on a disposable column. The resin was washed at 4°C with 2ml cold 10mM imidazole buffer until the flow-through reached the $O.D_{280}$ of 0.02-0.06. The resin was washed with 2ml cold 20mM imidazole buffer (300 mM NaCl, 50 mM phosphate buffer, 20 mM imidazole, pH 8) until the flow-through reached the $O.D_{280}$ of 0.02-0.06. The His-fusion protein was eluted by addition of 700μl cold 250mM imidazole buffer (300 mM NaCl, 50 mM phosphate buffer, 250 mM imidazole, pH 8) and fractions collected until the $O.D_{280}$ was 0.1. 21μl of each fraction were loaded on a 12% SDS gel.

His-fusion insoluble proteins large-scale purification.

A single colony was grown overnight at 37 °C on a LB + Amp agar plate. The bacteria were inoculated into 20 ml of LB+Amp liquid culture in a water bath shaker and grown overnight. Bacteria were diluted 1:30 into 600ml fresh medium and let to grow at the optimal temperature (37°C) to $O.D_{550}$ 0.6-0.8. Protein expression was induced by addition of 1 mM IPTG and the culture further incubated for three hours. The culture was centrifuged at 8000rpm at 4°C. The supernatant was discarded and the bacterial pellet was resuspended in 7.5 ml buffer B (urea 8M, 10mM Tris-HCl, 100mM phosphate buffer, pH 8.8). The cells were disrupted by sonication on ice for 30 sec at 40W using a Branson sonifier B-15, frozen and thawed twice and centrifuged again. The supernatant was stored at -20°C, while the pellets were resuspended in 2 ml guanidine buffer (6M guanidine hydrochloride, 100mM phosphate buffer, 10 mM Tris-HCl, pH 7.5) and treated in a homogenizer for 10 cycles. The product was centrifuged at 13000 rpm for 40 minutes. The supernatant was mixed with 150μl Ni^{2+}-resin (Pharmacia) (previously washed with buffer B) and incubated at room temperature with gentle agitation for 30 minutes. The sample was centrifuged at 700 g for 5 minutes at 4°C. The resin was washed twice with 10 ml buffer B for 10 minutes, resuspended in 1ml buffer B, and loaded on a disposable column. The resin was washed at room temperature with 2ml buffer B until the flow-through reached the OD_{280} of 0.02-0.06. The resin was washed with 2ml buffer C (urea 8M, 10mM Tris-HCl, 100mM phosphate buffer, pH 6.3) until the flow-through reached the $O.D_{280}$ of 0.02-0.06. The His-fusion protein was eluted by addition of 700μl elution buffer (urea 8M, 10mM Tris-HCl, 100mM phosphate buffer, pH 4.5) and fractions collected until the OD_{280} was 0.1. 21μl of each fraction were loaded on a 12% SDS gel.

Purification of His-fusion proteins.

For each clone to be purified as a His-fusion, a single colony was streaked out and grown overnight at 37°C on a LB/Amp (100 μg/ml) agar plate. An isolated colony from this plate was inoculated into 20ml of LB/Amp (100 μg/ml) liquid medium and grown overnight at 37°C with shaking. The overnight culture was diluted 1:30 into 600ml LB/Amp (100 μg/ml) liquid medium and allowed to grow at the optimal temperature (20-37°C) until the OD_{550nm} reached 0.6-0.8. Expression of recombinant protein was induced by addition of IPTG (final concentration 1.0mM) and the culture incubated for a further 3 hours. Bacteria were harvested by centrifugation at 8000xg for 15 min at 4°C.

The bacterial pellet was resuspended in 7.5ml of either (i) cold buffer A (300mM NaCl, 50mM phosphate buffer, 10mM imidazole, pH 8.0) for soluble proteins or (ii) buffer B (8M urea, 10mM Tris-HCl, 100mM phosphate buffer, pH 8.8) for insoluble proteins. Cells were disrupted by sonication on ice four times for 30 sec at 40W using a Branson sonifier 450 and centrifuged at 13 000xg for 30 min at 4°C. For insoluble proteins, pellets were resuspended in 2.0 ml buffer C (6M guanidine hydrochloride, 100mM phosphate buffer, 10mM Tris-HCl, pH 7.5) and treated with a Dounce homogenizer for 10 cycles. The homogenate was centrifuged at 13 000xg for 40 min and the supernatant retained.

Supernatants for both soluble and insoluble preparations were mixed with 150μl Ni^{2+}-resin (previously equilibrated with either buffer A or buffer B, as appropriate) and incubated at room temperature with gentle agitation for 30 min. The resin was Chelating Sepharose Fast Flow (Pharmacia), prepared according to manufacturers protocol. The batch-wise preparation was centrifuged at 700xg for 5 min at 4°C and the supernatant discarded. The resin was washed twice (batch-wise) with 10ml buffer A or B for 10 min, resuspended in 1.0 ml buffer A or B and loaded onto a disposable column. The resin continued to be washed with either (i) buffer A at 4°C or (ii) buffer B at room temperature, until the OD_{280nm} of the flow-through reached 0.02-0.01. The resin was further washed with either (i) cold buffer C (300mM NaCl, 50mM phosphate buffer, 20mM imidazole, pH 8.0) or (ii) buffer D (8M urea, 10mM Tris-HCl, 100mM phosphate buffer, pH 6.3) until the the OD_{280nm} of the flow-through reached 0.02-0.01. The His-fusion protein was eluted by addition of 700μl of either (i) cold elution buffer A (300mM NaCl, 50mM phosphate buffer, 250mM imidazole, pH 8.0) or (ii) elution buffer B (8 M urea, 10mM Tris-HCl, 100mM phosphate buffer, pH 4.5) and fractions

collected until the $O.D_{280nm}$ indicated all the recombinant protein was obtained. 20µl aliquots of each elution fraction were analyzed by SDS-PAGE using a 12% gel. Protein concentrations were estimated using the Bradford assay.

His-fusion proteins renaturation

In the cases where denaturation was required to solubilize proteins, a renaturation step was employed prior to immunization. Glycerol was added to the denatured fractions obtained above to give a final concentration of 10%(v/v). The proteins were diluted to 200µg/ml using dialysis buffer I (10% (v/v) glycerol, 0.5M arginine, 50mM phosphate buffer, 5.0mM reduced glutathione, 0.5mM oxidised glutathione, 2.0M urea, pH 8.8) and dialysed against the same buffer for 12-14 hours at 4°C. Further dialysis was performed with buffer II (10% (v/v) glycerol, 0.5M arginine, 50mM phosphate buffer, 5.0mM reduced glutathione, 0.5mM oxidised glutathione, pH 8.8) for 12-14 hours at 4°C.

Alternatively, 10% glycerol was added to the denatured proteins. The proteins were then diluted to 20µg/ml using dialysis buffer I (10% glycerol, 0.5M arginine, 50mM phosphate buffer, 5mM reduced glutathione, 0.5mM oxidised glutathione, 2M urea, pH 8.8) and dialysed against the same buffer at 4°C for 12-14 hours. The protein was further dialysed against dialysis buffer II (10% glycerol, 0.5M arginine, 50mM phosphate buffer, 5mM reduced glutathione, 0.5mM oxidised glutathione, pH 8.8) for 12-14 hours at 4°C.

Protein concentration was evaluated using the formula:

$$\text{Protein (mg/ml)} = (1.55 \times OD_{280}) - (0.76 \times OD_{260})$$

Purification of proteins

To analyse the solubility, pellets obtained from 3.0ml cultures were resuspended in 500µl buffer M1 (PBS pH 7.2). 25µl of lysozyme (10mg/ml) was added and the bacteria incubated for 15 min at 4°C. Cells were disrupted by sonication on ice four times for 30 sec at 40W using a Branson sonifier 450 and centrifuged at 13 000xg for 30 min at 4°C. The supernatant was collected and the pellet resuspended in buffer M2 [8M urea, 0.5M NaCl, 20mM imidazole and 0.1M NaH_2 PO_4] and incubated for 3 to 4 hours at 4°C. After centrifugation, the supernatant was collected and the pellet resuspended in buffer M3 [6M guanidinium-HCl, 0.5M NaCl, 20mM imidazole and 0.1M NaH_2PO_4] overnight at 4°C. The

supernatants from all steps were analysed by SDS-PAGE. Some proteins were found to be soluble in PBS, others need urea or guanidium-HCl for solubilization.

For preparative scale purifications, 500ml cultures were induced and fusion proteins solubilized in either buffer M1, M2 or M3 using the procedure described above. Crude extracts were loaded onto a Ni-NTA superflow column (Quiagen) equilibrated with buffer M1, M2 or M3 depending on the solubilization buffer employed. Unbound material was eluted by washing the column with the same buffer. The recombinant fusion protein was eluted with the corresponding buffer containing 500mM imidazole then dialysed against the same buffer in the absence of imidazole.

Mice immunisations

20μg of each purified protein are used to immunise mice intraperitoneally. In the case of some ORFs, Balb-C mice were immunised with $Al(OH)_3$ as adjuvant on days 1, 21 and 42, and immune response was monitored in samples taken on day 56. For other ORFs, CD1 mice could be immunised using the same protocol. For ORFs 25 and 40, CD1 mice were immunised using Freund's adjuvant, and the same immunisation protocol was used, except that the immune response was measured on day 42, rather than 56. Similarly, for still other ORFs, CD1 mice were immunised with Freund's adjuvant, but the immune response was measured on day 49. Alternatively, 20μg of each purified protein was mixed with Freund's adjuvant and used to immunise CD1 mice intraperitoneally. For many of the proteins, the immunization was performed on days 1, 21 and 35, and immune response was monitored in samples taken on days 34 and 49. For some proteins, the third immunization was performed on day 28, rather than 35, and the immune response was measured on days 20 and 42, rather than 34 and 49.

ELISA assay (sera analysis)

The acapsulated MenB M7 strain was plated on chocolate agar plates and incubated overnight at 37°C. Bacterial colonies were collected from the agar plates using a sterile dracon swab and inoculated into 7ml of Mueller-Hinton Broth (Difco) containing 0.25% Glucose. Bacterial growth was monitored every 30 minutes by following OD_{620}. The bacteria were let to grow until the OD reached the value of 0.3-0.4. The culture was centrifuged for 10 minutes at 10000 rpm. The supernatant was discarded and bacteria were washed once with PBS, resuspended in PBS containing 0.025% formaldehyde, and incubated

for 2 hours at room temperature and then overnight at 4°C with stirring. 100μl bacterial cells were added to each well of a 96 well Greiner plate and incubated overnight at 4°C. The wells were then washed three times with PBT washing buffer (0.1% Tween-20 in PBS). 200 μl of saturation buffer (2.7% Polyvinylpyrrolidone 10 in water) was added to each well and the plates incubated for 2 hours at 37°C. Wells were washed three times with PBT. 200 μl of diluted sera (Dilution buffer: 1% BSA, 0.1% Tween-20, 0.1% NaN_3 in PBS) were added to each well and the plates incubated for 90 minutes at 37°C. Wells were washed three times with PBT. 100 μl of HRP-conjugated rabbit anti-mouse (Dako) serum diluted 1:2000 in dilution buffer were added to each well and the plates were incubated for 90 minutes at 37°C. Wells were washed three times with PBT buffer. 100 μl of substrate buffer for HRP (25 ml of citrate buffer pH5, 10 mg of O-phenildiamine and 10 μl of H_2O) were added to each well and the plates were left at room temperature for 20 minutes. 100 μl H_2SO_4 was added to each well and OD_{490} was followed. The ELISA was considered positive when OD490 was 2.5 times the respective pre-immune sera.

Alternatively, The acapsulated MenB M7 strain was plated on chocolate agar plates and incubated overnight at 37°C. Bacterial colonies were collected from the agar plates using a sterile dracon swab and inoculated into Mueller-Hinton Broth (Difco) containing 0.25% Glucose. Bacterial growth was monitored every 30 minutes by following OD_{620}. The bacteria were let to grow until the OD reached the value of 0.3-0.4. The culture was centrifuged for 10 minutes at 10 000rpm. The supernatant was discarded and bacteria were washed once with PBS, resuspended in PBS containing 0.025% formaldehyde, and incubated for 1 hour at 37°C and then overnight at 4°C with stirring. 100μl bacterial cells were added to each well of a 96 well Greiner plate and incubated overnight at 4°C. The wells were then washed three times with PBT washing buffer (0.1% Tween-20 in PBS). 200μl of saturation buffer (2.7% Polyvinylpyrrolidone 10 in water) was added to each well and the plates incubated for 2 hours at 37°C. Wells were washed three times with PBT. 200μl of diluted sera (Dilution buffer: 1% BSA, 0.1% Tween-20, 0.1% NaN_3 in PBS) were added to each well and the plates incubated for 2 hours at 37°C. Wells were washed three times with PBT. 100μl of HRP-conjugated rabbit anti-mouse (Dako) serum diluted 1:2000 in dilution buffer were added to each well and the plates were incubated for 90 minutes at 37°C. Wells were washed three times with PBT buffer. 100μl of substrate buffer for HRP (25ml of citrate buffer pH5, 10mg of O-

phenildiamine and 10μl of H_2O_2) were added to each well and the plates were left at room temperature for 20 minutes. 100μl of 12.5% H_2SO_4 was added to each well and OD_{490} was followed. The ELISA titers were calculated abitrarely as the dilution of sera which gave an OD_{490} value of 0.4 above the level of preimmune sera. The ELISA was considered positive when the dilution of sera with OD_{490} of 0.4 was higher than 1:400.

FACScan bacteria Binding Assay procedure.

The acapsulated MenB M7 strain was plated on chocolate agar plates and incubated overnight at 37°C. Bacterial colonies were collected from the agar plates using a sterile dracon swab and inoculated into 4 tubes containing 8ml each Mueller-Hinton Broth (Difco) containing 0.25% glucose. Bacterial growth was monitored every 30 minutes by following OD_{620}. The bacteria were let to grow until the OD reached the value of 0.35-0.5. The culture was centrifuged for 10 minutes at 4000rpm. The supernatant was discarded and the pellet was resuspended in blocking buffer (1% BSA in PBS, 0.4% NaN_3) and centrifuged for 5 minutes at 4000rpm. Cells were resuspended in blocking buffer to reach OD_{620} of 0.07. 100μl bacterial cells were added to each well of a Costar 96 well plate. 100μl of diluted (1:100, 1:200, 1:400) sera (in blocking buffer) were added to each well and plates incubated for 2 hours at 4°C. Cells were centrifuged for 5 minutes at 4000rpm, the supernatant aspirated and cells washed by addition of 200μl/well of blocking buffer in each well. 100μl of R-Phicoerytrin conjugated $F(ab)_2$ goat anti-mouse, diluted 1:100, was added to each well and plates incubated for 1 hour at 4°C. Cells were spun down by centrifugation at 4000rpm for 5 minutes and washed by addition of 200μl/well of blocking buffer. The supernatant was aspirated and cells resuspended in 200μl/well of PBS, 0.25% formaldehyde. Samples were transferred to FACScan tubes and read. The condition for FACScan (Laser Power 15mW) setting were: FL2 on; FSC-H threshold:92; FSC PMT Voltage: E 01; SSC PMT: 474; Amp. Gains 6.1; FL-2 PMT: 586; compensation values: 0.

OMV preparations

Bacteria were grown overnight on 5 GC plates, harvested with a loop and resuspended in 10 ml 20mM Tris-HCl. Heat inactivation was performed at 56°C for 30 minutes and the bacteria disrupted by sonication for 10' on ice (50% duty cycle, 50% output). Unbroken cells were removed by centrifugation at 5000*g* for 10 minutes and the total cell envelope

fraction recovered by centrifugation at 50000*g* at 4°C for 75 minutes. To extract cytoplasmic membrane proteins from the crude outer membranes, the whole fraction was resuspended in 2% sarkosyl (Sigma) and incubated at room temperature for 20 minutes. The suspension was centrifuged at 10000*g* for 10 minutes to remove aggregates, and the supernatant further ultracentrifuged at 50000*g* for 75 minutes to pellet the outer membranes. The outer membranes were resuspended in 10mM Tris-HCl, pH8 and the protein concentration measured by the Bio-Rad Protein assay, using BSA as a standard.

Whole Extracts preparation

Bacteria were grown overnight on a GC plate, harvested with a loop and resuspended in 1ml of 20mM Tris-HCl. Heat inactivation was performed at 56°C for 30' minutes.

Western blotting

Purified proteins (500ng/lane), outer membrane vesicles (5μg) and total cell extracts (25μg) derived from MenB strain 2996 were loaded onto a 12% SDS-polyacrylamide gel and transferred to a nitrocellulose membrane. The transfer was performed for 2 hours at 150mA at 4°C, using transfer buffer (0.3 % Tris base, 1.44 % glycine, 20% (v/v) methanol). The membrane was saturated by overnight incubation at 4°C in saturation buffer (10% skimmed milk, 0.1% Triton X100 in PBS). The membrane was washed twice with washing buffer (3% skimmed milk, 0.1% Triton X100 in PBS) and incubated for 2 hours at 37°C with mice sera diluted 1:200 in washing buffer. The membrane was washed twice and incubated for 90 minutes with a 1:2000 dilution of horseradish peroxidase labelled anti-mouse Ig. The membrane was washed twice with 0.1% Triton X100 in PBS and developed with the Opti-4CN Substrate Kit (Bio-Rad). The reaction was stopped by adding water.

Bactericidal assay

MC58 and 2996 strains were grown overnight at 37°C on chocolate agar plates. 5-7 colonies were collected and used to inoculate 7ml Mueller-Hinton broth. The suspension was incubated at 37°C on a nutator and let to grow until OD_{620} was in between 0.5-0.8. The culture was aliquoted into sterile 1.5ml Eppendorf tubes and centrifuged for 20 minutes at maximum speed in a microfuge. The pellet was washed once in Gey's buffer (Gibco) and resuspended in the same buffer to an OD_{620} of 0.5, diluted 1:20000 in Gey's buffer and stored at 25°C.

50µl of Gey's buffer/1% BSA was added to each well of a 96-well tissue culture plate. 25µl of diluted (1:100) mice sera (dilution buffer: Gey's buffer/0.2% BSA) were added to each well and the plate incubated at 4°C. 25µl of the previously described bacterial suspension were added to each well. 25µl of either heat-inactivated (56°C waterbath for 30 minutes) or normal baby rabbit complement were added to each well. Immediately after the addition of the baby rabbit complement, 22µl of each sample/well were plated on Mueller-Hinton agar plates (time 0). The 96-well plate was incubated for 1 hour at 37°C with rotation and then 22µl of each sample/well were plated on Mueller-Hinton agar plates (time 1). After overnight incubation the colonies corresponding to time 0 and time 1h were counted.

Gene Variability

The *ORF4* and *919* genes were amplified by PCR on chromosomal DNA extracted from various Neisseria strains (see list of strains). The following oligonucleotides used as PCR primers were designed in the upstream and downstream regions of the genes:

orf 4.1	(forward)	CGAATCCGGACGGCAGGACTC
orf 4.3	(reverse)	GGCAGGGAATGGCGGATTAAAG
919.1	(forward)	AAAATGCCTCTCCACGGCTG or CTGCGCCCTGTGTTAAAATCCCCT
919.6	(reverse)	CAAATAAGAAAGGAATTTTG or GGTATCGCAAAACTTCGCCTTAATGCG

The PCR cycling conditions were:

1 cycle	2 min. at 94°
30 cycles	30 sec. at 94°
	30 sec. at ~ 54° or ~ 60° (in according to Tm of the primers)
	40 sec. at 72°
1 cycle	7 min. at 72°

The PCR products were purified from 1 % agarose gel and sequenced using the following primers:

orf 4.1 (forward)	CGAATCCGGACGGCAGGACTC
orf 4.2 (forward)	CGACCGCGCCTTTGGGACTG
orf 4.3 (reverse)	GGCAGGGAATGGCGGATTAAAG
orf 4.4 (reverse)	TCTTTGAGTTTGATCCAACC

919.1 (forward) AAAATGCCTCTCCACGGCTG or
CTGCGCCCTGTGTTAAAATCCCCT
919.2 (forward) ATCCTTCCGCCTCGGCTGCG
919.3 (forward) AAAACAGCGGCACAATCGAC
919.4 (forward) ATAAGGGCTACCTCAAACTC
919.5 (forward) GCGCGTGGATTATTTTTGGG
919.6 (reverse) CAAATAAGAAAGGAATTTTG or
GGTATCGCAAAACTTCGCCTTAATGCG
919.7 (reverse) CCCAAGGTAATGTAGTGCCG
919.8 (reverse) TAAAAAAAAGTTCGACAGGG
919.9 (reverse) CCGTCCGCCTGTCGTCGCCC
919.10 (reverse) TCGTTCCGGCGGGGTCGGGG

All documents cited herein are incorporated by reference in their entireties.

The following Examples are presented to illustrate, not limit, the invention

EXAMPLE 1

Using the above-described procedures, the following oligonucleotide primers were employed in the polymerase chain reaction (PCR) assay in order to clone the ORFs as indicated:

Table 1: Oligonucleotides used for PCR for Examples 2-10

ORF	Primer	Sequence	Restriction sites
279	Forward	CGCGGATCCCATATG-TTGCCTGCAATCACGATT <SEQ ID 3021>	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTAGAAGCGGGCGGCAA <SEQ ID 3022>	XhoI
519	Forward	CGCGGATCCCATATG-TTCAAATCCTTTGTCGTCA <SEQ ID 3023>	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTGGCGGTTTTGCTGC <SEQ ID 3024>	XhoI
576	Forward	CGCGGATCCCATATG-GCCGCCCCCGCATCT <SEQ ID 3025>	BamHI-NdeI
	Reverse	CCCGCTCGAG-ATTTACTTTTTTGATGTCGAC <SEQ ID 3026>	XhoI
919	Forward	CGCGGATCCCATATG-TGCCAAAGCAAGAGCATC <SEQ ID 3027>	BamHI-NdeI
	Reverse	CCCGCTCGAG-CGGGCGGTATTCGGG <SEQ ID 3028>	XhoI
121	Forward	CGCGGATCCCATATG-GAAACACAGCTTTACAT <SEQ ID 3029>	BamHI-NdeI

	Reverse	CCCGCTCGAG-ATAATAATATCCCGCGCCC <SEQ ID 3030>	XhoI
128	Forward	CGCGGATCCCATATG-ACTGACAACGCACT <SEQ ID 3031>	BamHI-NdeI
	Reverse	CCCGCTCGAG-GACCGCGTTGTCGAAA <SEQ ID 3032>	XhoI
206	Forward	CGCGGATCCCATATG-AAACACCGCCAACCGA <SEQ ID 3033>	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTCTGTAAAAAAAGTATGTGC <SEQ ID 3034>	XhoI
287	Forward	CCGGAATTCTAGCTAGC-CTTTCAGCCTGCGGG <SEQ ID 3035>	EcoRI-NheI
	Reverse	CCCGCTCGAG-ATCCTGCTCTTTTTTGCC <SEQ ID 3036>	XhoI
406	Forward	CGCGGATCCCATATG-TGCGGGACACTGACAG <SEQ ID 3037>	BamHI-NdeI
	Reverse	CCCGCTCGAG-AGGTTGTCCTTGTCTATG <SEQ ID 3038>	XhoI

Localization of the ORFs

The following DNA and amino acid sequences are identified by titles of the following form: [g, m, or a] [#].[seq or pep], where "g" means a sequence from *N. gonorrhoeae*, "m" means a sequence from *N. meningitidis B*, and "a" means a sequence from *N. meningitidis A*; "#" means the number of the sequence; "seq" means a DNA sequence, and "pep" means an amino acid sequence. For example, "g001.seq" refers to an *N. gonorrohoeae* DNA sequence, number 1. The presence of the suffix "-1" to these sequences indicates an additional sequence found for the same ORF, thus, data for an ORF having both an unsuffixed and a suffixed sequence designation applies to both such designated sequences. Further, open reading frames are identified as ORF #, where "#" means the number of the ORF, corresponding to the number of the sequence which encodes the ORF, and the ORF designations may be suffixed with ".ng" or ".a", indicating that the ORF corresponds to a *N. gonorrhoeae* sequence or a *N. meningitidis A* sequence, respectively. The word "partial" before a sequence indicates that the sequence may be a partial or a complete ORF. Computer analysis was performed for the comparisons that follow between "g", "m", and "a" peptide sequences; and therein the "pep" suffix is implied where not expressly stated. Further, in the event of a conflict between the text immediately preceding and describing which sequences are being compared, and the

designated sequences being compared, the designated sequence controls and is the actual sequence being compared.

ORF: contig:

279 gnm4.seq

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3039>:
m279.seq

```
  1  ATAACGCGGA TTTGCGGCTG CTTGATTTCA ACGGTTTTCA GGGCTTCGGC
 51  AAGTTTGTCG GCGGCGGGTT TCATCAGGCT GCAATGGGAA GGTACGGACA
101  CGGGCAGCGG CAGGGCGCGT TTGGCACCGG CTTCTTTGGC GGCAGCCATG
151  GCGCGTCCGA CGGCGGCGGC GTTGCCTGCA ATCACGATTT GTCCGGGTGA
201  GTTGAAGTTG ACGGCTTCGA CCACTTCGCT TTGGGCGGCT TCGGCACAAA
251  TGGCTTTAAC CTGCTCATCT TCCAAGCCGA G[illegible]TCGCCGC CATTGCGCCC
301  ACGCCTTGCG GTACGGCGGA CTGCATCAGT TCGGCGCGCA GGCGCACGAG
351  TTTGACCGCG TCGGCAAAAT TCAATGCGCC GGCGGCA[illegible]AGTGCGGTGT
401  ATTCGCCGAG GCTGTGTCCG GCAACGGCGG CAGGCGTT[illegible]T GCCGCCCGCT
451  TCTAAATAG
```

This corresponds to the amino acid sequence <SEQ ID 3040; ORF 279>:
m279.pep

```
  1  ITRICGCLIS TVFRASASLS AAGFIRLQWE GTDTGSGRAR LAPASLAAAM
 51  ARPTAAALPA ITICPGELKL TASTTSLWAA SAQMALTCSS SKPRIAAIAP
101  TPCGTADCIS SARRRTSLTA SAKFNAPAAT SAVYSPRLCP ATAAGVLPPA
151  SK*
```

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 3041>:
g279.seq

```
  1  atgacgcgga tttgcggctg cttgatttca acggttttga gtgtttcggc
 51  aagtttgtcg gcggcgggtt tcatcaggct gcaatgggaa ggaacggata
101  ccggcagcgg cagggcgcgt ttggctccgg cttctttggc ggcagccatg
151  gtgcgtccga cggcggcggc gttgcctgca atcacgactt gtccgggcga
201  gttgaagttg acggcttcga ccacttcgcc ctgtgcggat tcggcacaaa
251  tctgcctgac ctgttcatct tccaaaccca aaatggccgc cattgcgcct
301  acgccttgcg gtacggcgga ctgcatcagt tcggcgcgca ggcggacgag
351  tttgacggca tcggcaaaat ccaatgcttc ggcggcgaca agcgcggtgt
401  attcgccgag gctgtgtccg gcaacggcgg caggcgtttt gccgcccact
451  tccaaatag
```

This corresponds to the amino acid sequence <SEQ ID 3042; ORF 279.ng>:
g279.pep

```
  1  MTRICGCLIS TVLSVSASLS AAGFIRLQWE GTDTGSGRAR LAPASLAAAM
 51  VRPTAAALPA ITTCPGELKL TASTTSPCAD SAQICLTCSS SKPKMAAIAP
101  TPCGTADCIS SARRRTSLTA SAKSNASAAT SAVYSPRLCP ATAAGVLPPT
151  SK*
```

ORF 279 shows 89.5% identity over a 152 aa overlap with a predicted ORF (ORF 279.ng) from *N. gonorrhoeae:*

```
                      10        20        30        40        50        60
m279.pep      ITRICGCLISTVFRASASLSAAGFIRLQWEGTDTGSGRARLAPASLAAAMARPTAAALPA
              :||||||||||:  :||||||||||||||||||||||||||||||||||:|||||||||
g279          MTRICGCLISTVLSVSASLSAAGFIRLQWEGTDTGSGRARLAPASLAAAMVRPTAAALPA
                      10        20        30        40        50        60
```

```
                   70        80        90       100       110       120
m279.pep   ITICPGELKLTASTTSLWAASAQMALTCSSSKPRIAAIAPTPCGTADCISSARRRTSLTA
           || |||||||||||||  | |||: ||||||||::|||||||||||||||||||||||||
g279       ITTCPGELKLTASTTSPCADSAQICLTCSSSKPKMAAIAPTPCGTADCISSARRRTSLTA
                   70        80        90       100       110       120

                  130       140       150
m279.pep   SAKFNAPAATSAVYSPRLCPATAAGVLPPASKX
           ||| || ||||||||||||||||||||||:|||
g279       SAKSNASAATSAVYSPRLCPATAAGVLPPTSKX
                  130       140       150
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3043>:

```
a279.seq
      1  ATGACNCNGA TTTGCGGCTG CTTGATTTCA ACGGTTTNNA GGGCTTCGGC
     51  GAGTTTGTCG GCGGCGGGTT TCATGAGGCT GCAATGGGAA GGTACNGACA
    101  CNGGCAGCGG CAGGGCGCGT TTGGCGCCGG CTTCTTTGGC GGCAAGCATA
    151  GCGCGCTCGA CGGCGGCGGC ATTGCCTGCA ATCACGACTT GTCCGGGCGA
    201  GTTGAAGTTG ACGGCTTCAA CCACTTCATC CTGTGCGGAT TCGGCGCAAA
    251  TTTGTTTTAC CTGTTCATCT TCCAAGCCGA GAATCGCCGC CATTGCGCCC
    301  ACGCCTTGCG GTACGGCGGA CTGCATCAGT TCGGCGCGCA NGCGCACGAG
    351  TTTGACCGCG TCGGCAAAAT CCAATGCGCC GGCGGCAACN AGTGCGGTGT
    401  ATTCGCCGAN GCTGTGTCCG GCAACGGCGG CAGGCGTTTT GCCGCCCGCT
    451  TCCGAATAG
```

This corresponds to the amino acid sequence <SEQ ID 3044; ORF 279.a>:

```
a279.pep
      1  MTXICGCLIS TVXRASASLS AAGFMRLQWE GTDTGSGRAR LAPASLAASI
     51  ARSTAAALPA ITTCPGELKL TASTTSSCAD SAQICFTCSS SKPRIAAIAP
    101  TPCGTADCIS SARXRTSLTA SAKSNAPAAT SAVYSPXLCP ATAAGVLPPA
    151  SE*
```

m279/a279 ORFs 279 and 279.a showed a 88.2% identity in 152 aa overlap

```
                   10        20        30        40        50        60
m279.pep   ITRICGCLISTVFRASASLSAAGFIRLQWEGTDTGSGRARLAPASLAAAMARPTAAALPA
           :| ||||||||| |||||||||||:|||||||||||||||||||||||::|| |||||||
a279       MTXICGCLISTVXRASASLSAAGFMRLQWEGTDTGSGRARLAPASLAASIARSTAAALPA
                   10        20        30        40        50        60

                   70        80        90       100       110       120
m279.pep   ITICPGELKLTASTTSLWAASAQMALTCSSSKPRIAAIAPTPCGTADCISSARRRTSLTA
           || |||||||||||||  | |||: :||||||||||||||||||||||||||| ||||||
a279       ITTCPGELKLTASTTSSCADSAQICFTCSSSKPRIAAIAPTPCGTADCISSARXRTSLTA
                   70        80        90       100       110       120

                  130       140       150
m279.pep   SAKFNAPAATSAVYSPRLCPATAAGVLPPASKX
           ||| |||||||||||| ||||||||||||||:|
a279       SAKSNAPAATSAVYSPXLCPATAAGVLPPASEX
                  130       140       150
```

519 and 519-1 gnm7.seq

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3045>:

```
m519.seq  (partial)
      1  ..TCCGTTATCG GGCGTATGGA GTTGGACAAA ACGTTTGAAG AACGCGACGA
     51    AATCAACAGT ACTGTTGTTG CGGCTTTGGA CGAGGCGGCC GGGgCTTgGG
    101    GTGTGAAGGT TTTGCGTTAT GAGATTAAAG ACTTGGTTCC GCCGCAAGAA
    151    ATCCTTCGCT CAATGCAGGC GCAAATTACT GCCGAACGCG AAAAACGCGC
    201    CCGTATCGCC GAATCCGAAG GTCGTAAAAT CGAACAAATC AACCTTGCCA
```

```
251    GTGGTCAGCG CGAAGCCGAA ATCCAACAAT CCGAAGGCGA GGCTCAGGCT
301    GCGGTCAATG CGTCAAATGC CGAGAAAATC GCCCGCATCA ACCGCGCCAA
351    AGGTGAAGCG GAATCCTTGC GCCTTGTTGC CGAAGCCAAT GCCGAAGCCA
401    TCCGTCAAAT TGCCGCCGCC CTTCAAACCC AAGGCGGTGC GGATGCGGTC
451    AATCTGAAGA TTGCGGAACA ATACGTCGCT GCGTTCAACA ATCTTGCCAA
501    AGAAAGCAAT ACGCTGATTA TGCCCGCCAA TGTTGCCGAC ATCGGCAGCC
551    TGATTTCTGC CGGTATGAAA ATTATCGACA GCAGCAAAAC CGCCAAaTAA
```

This corresponds to the amino acid sequence <SEQ ID 3046; ORF 519>:

```
m519.pep     (partial)
      1    ..SVIGRMELDK TFEERDEINS TVVAALDEAA GAWGVKVLRY EIKDLVPPQE
     51      ILRSMQAQIT AEREKRARIA ESEGRKIEQI NLASGQREAE IQQSEGEAQA
    101      AVNASNAEKI ARINRAKGEA ESLRLVAEAN AEAIRQIAAA LQTQGGADAV
    151      NLKIAEQYVA AFNNLAKESN TLIMPANVAD IGSLISAGMK IIDSSKTAK*
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3047>:

```
g519.seq
      1  atggaatttt tcattatctt gttggcagcc gtggccgttt tcggcttcaa
     51  atcctttgtc gtcatccccc agcaggaagt ccacgttgtc gaaaggctcg
    101  ggcgtttcca tcgcgccctg acggccggtt tgaatatttt gattcccttt
    151  atcgaccgcg tcgcctaccg ccattcgctg aaagaaatcc ctttagacgt
    201  acccagccag gtctgcatca cgcgcgataa tacgcaattg actgttgacg
    251  gcatcatcta tttccaagta accgatccca aactcgcctc atacggttcg
    301  agcaactaca ttatggcaat tacccagctt gcccaaacga cgctgcgttc
    351  cgttatcggg cgtatggagt tggacaaaac gtttgaagaa cgcgacgaaa
    401  tcaacagtac cgtcgtctcc gccctcgatg aagccgccgg ggcttggggt
    451  gtgaaagtcc tccgttacga aatcaaggat ttggttccgc cgcaagaaat
    501  ccttcgcgca atgcaggcac aaattaccgc cgaacgcgaa aaacgcgccc
    551  gtattgccga atccgaaggc cgtaaaatcg aacaaatcaa ccttgccagt
    601  ggtcagcgtg aagccgaaat ccaacaatcc gaaggcgagg ctcaggctgc
    651  ggtcaatgcg tccaatgccg agaaaatcgc ccgcatcaac cgcgccaaag
    701  gcgaagcgga atccctgcgc cttgttgccg aagccaatgc cgaagccaac
    751  cgtcaaattg ccgccgccct tcaaacccaa agcggggcgg atgcggtcaa
    801  tctgaagatt gcgggacaat acgttaccgc gttcaaaaat cttgccaaag
    851  aagacaatac gcggattaag cccgccaagg ttgccgaaat cgggaaccct
    901  aatttcggc ggcatgaaaa attttcgcca gaagcaaaaa cggccaaata
    951  a
```

This corresponds to the amino acid sequence <SEQ ID 3048; ORF 519.ng>:

```
g519.pep
      1  MEFFIILLAA VAVFGFKSFV VIPQQEVHVV ERLGRFHRAL TAGLNILIPF
     51  IDRVAYRHSL KEIPLDVPSQ VCITRDNTQL TVDGIIYFQV TDPKLASYGS
    101  SNYIMAITQL AQTTLRSVIG RMELDKTFEE RDEINSTVVS ALDEAAGAWG
    151  VKVLRYEIKD LVPPQEILRA MQAQITAERE KRARIAESEG RKIEQINLAS
    201  GQREAEIQQS EGEAQAAVNA SNAEKIARIN RAKGEAESLR LVAEANAEAN
    251  RQIAAALQTQ SGADAVNLKI AGQYVTAFKN LAKEDNTRIK PAKVAEIGNP
    301  NFRRHEKFSP EAKTAK*
```

ORF 519 shows 87.5% identity over a 200 aa overlap with a predicted ORF (ORF 519.ng) from *N. gonorrhoeae:*

```
m519/g519

                                                       10        20        30
m519.pep                                     SVIGRMELDKTFEERDEINSTVVAALDEAA
                                             ||||||||||||||||||||||||:|||||
g519        YFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIGRMELDKTFEERDEINSTVVSALDEAA
              90       100       110       120       130       140

                    40        50        60        70        80        90
m519.pep    GAWGVKVLRYEIKDLVPPQEILRSMQAQITAEREKRARIAESEGRKIEQINLASGQREAE
```

```
                ||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
g519            GAWGVKVLRYEIKDLVPPQEILRAMQAQITAEREKRARIAESEGRKIEQINLASGQREAE
                  150       160       170       180       190       200

                        100       110       120       130       140       150
m519.pep        IQQSEGEAQAAVNASNAEKIARINRAKGEAESLRLVAEANAEAIRQIAAALQTQGGADAV
                ||||||||||||||||||||||||||||||||||||||||||| ||||||||||:|||||
g519            IQQSEGEAQAAVNASNAEKIARINRAKGEAESLRLVAEANAEANRQIAAALQTQSGADAV
                  210       220       230       240       250       260

                        160       170       180       190       200
m519.pep        NLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL-ISAGMKIIDSSKTAK
                ||||| |||:||:|||||:|| | ||:||:||:  :    |:   :||||
g519            NLKIAGQYVTAFKNLAKEDNTRIKPAKVAEIGNPNFRRHEKFSPEAKTAK
                  270       280       290       300       310
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3049>:

```
a519.seq
      1  ATGGAATTTT TCATTATCTT GCTGGCAGCC GTCGTTGTTT TCGGCTTCAA
     51  ATCCTTTGTT GTCATCCCAC AGCAGGAAGT CCACGTTGTC GAAAGGCTCG
    101  GGCGTTTCCA TCGCGCCCTG ACGGCCGGTT TGAATATTTT GATTCCCTTT
    151  ATCGACCGCG TCGCCTACCG CCATTCGCTG AAAGAAATCC CTTTAGACGT
    201  ACCCAGCCAG GTCTGCATCA CGCGCGACAA TACGCAGCTG ACTGTTGACG
    251  GTATCATCTA TTTCCAAGTA ACCGACCCCA AACTCGCCTC ATACGGTTCG
    301  AGCAACTACA TTATGGCGAT TACCCAGCTT GCCCAAACGA CGCTGCGTTC
    351  CGTTATCGGG CGTATGGAAT TGGACAAAAC GTTTGAAGAA CGCGACGAAA
    401  TCAACAGCAC CGTCGTCTCC GCCCTCGATG AAGCCGCCGG AGCTTGGGGT
    451  GTGAAGGTTT TGCGTTATGA GATTAAAGAC TTGGTTCCGC CGCAAGAAAT
    501  CCTTCGCTCA ATGCAGGCGC AAATTACTGC TGAACGCGAA AAACGCGCCC
    551  GTATCGCCGA ATCCGAAGGT CGTAAAATCG AACAAATCAA CCTTGCCAGT
    601  GGTCAGCGCG AAGCCGAAAT CCAACAATCC GAAGGCGAGG CTCAGGCTGC
    651  GGTCAATGCG TCAAATGCCG AGAAAATCGC CCGCATCAAC CGCGCCAAAG
    701  GTGAAGCGGA ATCCTTGCGC CTTGTTGCCG AAGCCAATGC CGAAGCCATC
    751  CGTCAAATTG CCGCCGCCCT TCAAACCCAA GGCGGTGCGG ATGCGGTCAA
    801  TCTGAAGATT GCGGAACAAT ACGTCGCCGC GTTCAACAAT CTTGCCAAAG
    851  AAAGCAATAC GCTGATTATG CCCGCCAATG TTGCCGACAT CGGCAGCCTG
    901  ATTTCTGCCG GTATGAAAAT TATCGACAGC AGCAAAACCG CCAAATAA
```

This corresponds to the amino acid sequence <SEQ ID 3050; ORF 519.a>:

```
a519.pep
      1  MEFFIILLAA VVVFGFKSFV VIPQQEVHVV ERLGRFHRAL TAGLNILIPF
     51  IDRVAYRHSL KEIPLDVPSQ VCITRDNTQL TVDGIIYFQV TDPKLASYGS
    101  SNYIMAITQL AQTTLRSVIG RMELDKTFEE RDEINSTVVS ALDEAAGAWG
    151  VKVLRYEIKD LVPPQEILRS MQAQITAERE KRARIAESEG RKIEQINLAS
    201  GQREAEIQQS EGEAQAAVNA SNAEKIARIN RAKGEAESLR LVAEANAEAI
    251  RQIAAALQTQ GGADAVNLKI AEQYVAAFNN LAKESNTLIM PANVADIGSL
    301  ISAGMKIIDS SKTAK*
```

m519/a519 ORFs 519 and 519.a showed a 99.5% identity in 199 aa overlap

```
                                                      10        20        30
m519.pep                                      SVIGRMELDKTFEERDEINSTVVAALDEAA
                                              |||||||||||||||||||||||:||||||
a519            YFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIGRMELDKTFEERDEINSTVVSALDEAA
                   90       100       110       120       130       140

                          40        50        60        70        80        90
m519.pep        GAWGVKVLRYEIKDLVPPQEILRSMQAQITAEREKRARIAESEGRKIEQINLASGQREAE
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a519            GAWGVKVLRYEIKDLVPPQEILRSMQAQITAEREKRARIAESEGRKIEQINLASGQREAE
                  150       160       170       180       190       200

                        100       110       120       130       140       150
```

```
m519.pep    IQQSEGEAQAAVNASNAEKIARINRAKGEAESLRLVAEANAEAIRQIAAALQTQGGADAV
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a519        IQQSEGEAQAAVNASNAEKIARINRAKGEAESLRLVAEANAEAIRQIAAALQTQGGADAV
             210       220       230       240       250       260

                   160       170       180       190       200
m519.pep    NLKIAEQYVAAFNNLAKESNTLIMPANVADIGSLISAGMKIIDSSKTAKX
            ||||||||||||||||||||||||||||||||||||||||||||||||||
a519        NLKIAEQYVAAFNNLAKESNTLIMPANVADIGSLISAGMKIIDSSKTAKX
             270       280       290       300       310
```

Further work revealed the DNA sequence identified in *N. meningitidis* <SEQ ID 3051>:

```
m519-1.seq
      1  ATGGAATTTT TCATTATCTT GTTGGTAGCC GTCGCCGTTT TCGGTTTCAA
     51  ATCCTTTGTT GTCATCCCAC AACAGGAAGT CCACGTTGTC GAAAGGCTGG
    101  GGCGTTTCCA TCGCGCCCTG ACGGcCGGTT TGAATATTTT GATTCCCTTT
    151  ATCGACCGCG TCGCCTACCG CCATTCGCTG AAAGAAATCC CTTTAGACGT
    201  ACCCAGCCAG GTCTGCATCA CGCGCGACAA TACGCAGCTG ACTGTTGACG
    251  GCATCATCTA TTTCCAAGTA ACCGACCCCA AACTCGCCTC ATACGGTTCG
    301  AGCAACTACA TTATGGCGAT TACCCAGCTT GCCCAAACGA CGCTGCGTTC
    351  CGTTATCGGG CGTATGGAGT TGGACAAAAC GTTTGAAGAA CGCGACGAAA
    401  TCAACAGTAC TGTTGTTGCG GCTTTGGACG AGCCCGCCGG GGCTTGGGGT
    451  GTGAAGGTTT TGCGTTATGA GATTAAAGAC TTGGTTCCGC CGCAAGAAAT
    501  CCTTCGCTCA ATGCAGGCGC AAATTACTGC CGAACGCGAA AAACGCGCCC
    551  GTATCGCCGA ATCCGAAGGT CGTAAAATCG AACAAATCAA CCTTGCCAGT
    601  GGTCAGCGCG AAGCCGAAAT CCAACAATCC GAAGGCGAGG CTCAGGCTGC
    651  GGTCAATGCG TCAAATGCCG AGAAAATCGC CCGCATCAAC CGCGCCAAAG
    701  GTGAAGCGGA ATCCTTGCGC CTTGTTGCCG AAGCCAATGC CGAAGCCATC
    751  CGTCAAATTG CCGCCGCCCT TCAAACCCAA GGCGGTGCGG ATGCGGTCAA
    801  TCTGAAGATT GCGGAACAAT ACGTCGCTGC GTTCAACAAT CTTGCCAAAG
    851  AAAGCAATAC GCTGATTATG CCCGCCAATG TTGCCGACAT CGGCAGCCTG
    901  ATTTCTGCCG GTATGAAAAT TATCGACAGC AGCAAAACCG CCAAATAA
```

This corresponds to the amino acid sequence <SEQ ID 3052; ORF 519-1>:

```
m519-1.
      1  MEFFIILLVA VAVFGFKSFV VIPQQEVHVV ERLGRFHRAL TAGLNILIPF
     51  IDRVAYRHSL KEIPLDVPSQ VCITRDNTQL TVDGIIYFQV TDPKLASYGS
    101  SNYIMAITQL AQTTLRSVIG RMELDKTFEE RDEINSTVVA ALDEAAGAWG
    151  VKVLRYEIKD LVPPQEILRS MQAQITAERE KRARIAESEG RKIEQINLAS
    201  GQREAEIQQS EGEAQAAVNA SNAEKIARIN RAKGEAESLR LVAEANAEAI
    251  RQIAAALQTQ GGADAVNLKI AEQYVAAFNN LAKESNTLIM PANVADIGSL
    301  ISAGMKIIDS SKTAK*
```

The following DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3053>:

```
g519-1.seq
      1  ATGGAATTTT TCATTATCTT GTTGGCAGCC GTCGCCGTTT TCGGCTTCAA
     51  ATCCTTTGTC GTCATCCCCC AGCAGGAAGT CCACGTTGTC GAAAGGCTCG
    101  GGCGTTTCCA TCGCGCCCTG ACGGCCGGTT TGAATATTTT GATTCCCTTT
    151  ATCGACCGCG TCGCCTACCG CCATTCGCTG AAAGAAATCC CTTTAGACGT
    201  ACCCAGCCAG GTCTGCATCA CGCGCGATAA TACGCAATTG ACTGTTGACG
    251  GCATCATCTA TTTCCAAGTA ACCGATCCCA AACTCGCCTC ATACGGTTCG
    301  AGCAACTACA TTATGGCAAT TACCCAGCTT GCCCAAACGA CGCTGCGTTC
    351  CGTTATCGGG CGTATGGAGT TGGACAAAAC GTTTGAAGAA CGCGACGAAA
    401  TCAACAGTAC CGTCGTCTCC GCCCTCGATG AAGCCGCCGG GGCTTGGGGT
    451  GTGAAAGTCC TCCGTTACGA AATCAAGGAT TTGGTTCCGC CGCAAGAAAT
    501  CCTTCGCGCA ATGCAGGCAC AAATTACCGC CGAACGCGAA AAACGCGCCC
    551  GTATTGCCGA ATCCGAAGGC CGTAAAATCG AACAAATCAA CCTTGCCAGT
    601  GGTCAGCGTG AAGCCGAAAT CCAACAATCC GAAGGCGAGG CTCAGGCTGC
    651  GGTCAATGCG TCCAATGCCG AGAAAATCGC CCGCATCAAC CGCGCCAAAG
    701  GCGAAGCGGA ATCCCTGCGC CTTGTTGCCG AAGCCAATGC CGAAGCCATC
    751  CGTCAAATTG CCGCCGCCCT TCAAACCCAA GGCGGGGCGG ATGCGGTCAA
    801  TCTGAAGATT GCGGAACAAT ACGTAGCCGC GTTCAACAAT CTTGCCAAAG
```

```
851  AAAGCAATAC GCTGATTATG CCCGCCAATG TTGCCGACAT CGGCAGCCTG
901  ATTTCTGCCG GCATGAA[illegible]T TATCGACAGC AGCAAAACCG CCAAATAA
```

This corresponds to the amino acid sequence <SEQ ID 3054; ORF 519-1.ng>:

```
g519-1.pep
   1  MEFFIILLAA VAVFGFKSFV VIPQQEVHVV ERLGRFHRAL TAGLNILIPF
  51  IDRVAYRHSL KEIPLDVPSQ VCITRDNTQL TVDGIIYFQV TDPKLASYGS
 101  SNYIMAITQL AQTTLRSVIG RMELDKTFEE RDEINSTVVS ALDEAAGAWG
 151  VKVLRYEIKD LVPPQEILRA MQAQITAERE KRARIAESEG RKIEQINLAS
 201  GQREAEIQQS EGEAQAAVNA SNAEKIARIN RAKGEAESLR LVAEANAEAI
 251  RQIAAALQTQ GGADAVNLKI AEQYVAAFNN LAKESNTLIM PANVADIGSL
 301  ISAGMKIIDS SKTAK*
```

m519-1/g519-1 ORFs 519-1 and 519-1.ng showed a 99.0% identity in 315 aa overlap

```
                    10        20        30        40        50        60
g519-1.pep  MEFFIILLAAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
            |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
m519-1      MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
                    10        20        30        40        50        60

                    70        80        90       100       110       120
g519-1.pep  KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m519-1      KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
                    70        80        90       100       110       120

                   130       140       150       160       170       180
g519-1.pep  RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRAMQAQITAERE
            |||||||||||||||||||||||:||||||||||||||||||||||||:|||||||||||
m519-1      RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
                   130       140       150       160       170       180

                   190       200       210       220       230       240
g519-1.pep  KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m519-1      KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
                   190       200       210       220       230       240

                   250       260       270       280       290       300
g519-1.pep  LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m519-1      LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
                   250       260       270       280       290       300

                   310
g519-1.pep  ISAGMKIIDSSKTAKX
            ||||||||||||||||
m519-1      ISAGMKIIDSSKTAKX
                   310
```

The following DNA sequence was identified in *N. meningitidis* <SEQ ID 3055>:

```
a519-1.seq
   1  ATGGAATTTT TCATTATCTT GCTGGCAGCC GTCGTTGTTT TCGGCTTCAA
  51  ATCCTTTGTT GTCATCCCAC AGCAGGAAGT CCACGTTGTC GAAAGGCTCG
 101  GGCGTTTCCA TCGCGCCCTG ACGGCCGGTT TGAATATTTT GATTCCCTTT
 151  ATCGACCGCG TCGCCTACCG CCATTCGCTG AAAGAAATCC CTTTAGACGT
 201  ACCCAGCCAG GTCTGCATCA CGCGCGACAA TACGCAGCTG ACTGTTGACG
 251  GTATCATCTA TTTCCAAGTA ACCGACCCCA AACTCGCCTC ATACGGTTCG
 301  AGCAACTACA TTATGGCGAT TACCCAGCTT GCCCAAACGA CGCTGCGTTC
 351  CGTTATCGGG CGTATGGAAT TGGACAAAAC GTTTGAAGAA CGCGACGAAA
```

```
401  TCAACAGCAC CGTCGTCTCC GCCCTCGATG AAGCCGCCGG AGCTTGGGGT
451  GTGAAGGTTT TGCGTTATGA GATTAAAGAC TTGGTTCCGC CGCAAGAAAT
501  CCTTCGCTCA ATGCAGGCGC AAATTACTGC TGAACGCGAA AAACGCGCCC
551  GTATCGCCGA ATCCGAAGGT CGTAAAATCG AACAAATCAA CCTTGCCAGT
601  GGTCAGCGCG AAGCCGAAAT CCAACAATCC GAAGGCGAGG CTCAGGCTGC
651  GGTCAATGCG TCAAATGCCG AGAAAATCGC CCGCATCAAC CGCGCCAAAG
701  GTGAAGCGGA ATCCTTGCGC CTTGTTGCCG AAGCCAATGC CGAAGCCATC
751  CGTCAAATTG CCGCCGCCCT TCAAACCCAA GGCGGTGCGG ATGCGGTCAA
801  TCTGAAGATT GCGGAACAAT ACGTCGCCGC GTTCAACAAT CTTGCCAAAG
851  AAAGCAATAC GCTGATTATG CCCGCCAATG TTGCCGACAT CGGCAGCCTG
901  ATTTCTGCCG GTATGAAAAT TATCGACAGC AGCAAAACCG CCAAATAA
```

This corresponds to the amino acid sequence <SEQ ID 3056; ORF 519-1.a>:

```
a519-1.pep.
  1  MEFFIILLAA VVVFGFKSFV VIPQQEVHVV ERLGRFHRAL TAGLNILIPF
 51  IDRVAYRHSL KEIPLDVPSQ VCITRDNTQL TVDGIIYFQV TDPKLASYGS
101  SNYIMAITQL AQTTLRSVIG RMELDKTFEE RDEINSTVVS ALDEAAGAWG
151  VKVLRYEIKD LVPPQEILRS MQAQITAERE KRARIAESEG RKIEQINLAS
201  GQREAEIQQS EGEAQAAVNA SNAEKIARIN RAKGEAESLR LVAEANAEAI
251  RQIAAALQTQ GGADAVNLKI AEQYVAAFNN LAKESNTLIM PANVADIGSL
301  ISAGMKIIDS SKTAK*
```

m519-1/a519-1 ORFs 519-1 and 519-1.a showed a 99.0% identity in 315 aa overlap

```
                      10        20        30        40        50        60
a519-1.pep    MEFFIILLAAVVVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
              ||||||||:||:||||||||||||||||||||||||||||||||||||||||||||||||
m519-1        MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
                      10        20        30        40        50        60

                      70        80        90       100       110       120
a519-1.pep    KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m519-1        KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
                      70        80        90       100       110       120

                     130       140       150       160       170       180
a519-1.pep    RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
              |||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
m519-1        RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
                     130       140       150       160       170       180

                     190       200       210       220       230       240
a519-1.pep    KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m519-1        KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
                     190       200       210       220       230       240

                     250       260       270       280       290       300
a519-1.pep    LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m519-1        LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
                     250       260       270       280       290       300

                     310
a519-1.pep    ISAGMKIIDSSKTAKX
              |||||||||||||||
m519-1        ISAGMKIIDSSKTAKX
                     310
```

576 and 576-1 gnm22.seq

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3057>:

```
m576.seq.. (partial)
  1  ..ATGCAGCAGG CAAGCTATGC GATGGGCGTG GACATCGGAC GCTCCCTGAA
 51    GCAAATGAAG GAACAGGGCG CGGAAATCGA TTTGAAAGTC TTTACCGAAG
101    CCATGCAGGC AGTGTATGAC GGCAAAGAAA TCAAAATGAC CGAAGAGCAG
151    GCTCAGGAAG TCATGATGAA ATTCCTTCAG GAACAACAGG CTAAAGCCGT
201    AGAAAAACAC AAGGCGGACG CGAAGGCCAA TAAAGAAAAA GGCGAAGCCT
251    TTCTGAAAGA AAATGCCGCC AAAGACGGCG TGAAGACCAC TGCTTCCGGC
301    CTGCAATACA AAATCACCAA ACAGGGCGAA GGCAAACAGC CGACCAAAGA
351    CGACATCGTT ACCGTGGAAT ACGAAGGCCG CCTGATTGAC GGTACGGTAT
401    TCGACAGCAG CAAAGCCAAC GGCGGCCCGG TCACCTTCCC TTTGAGCCAA
451    GTGATTCCGG GTTGGACCGA AGgCGTACAG CTTCTGAAAG AAGGCGGCGA
501    AGCCACGTTC TACATCCCGT CCAACCTTGC CTACCGCGAA CAGGGTGCGG
551    GCGACAAAAT CGGTCCGAAC GCCACTTTGG TATTTGATGT GAAACTGGTC
601    AAAATCGGCG CACCCGAAAA CGCGCCCGCC AAGCAGCCGG CTCAAGTCGA
651    CATCAAAAAA GTAAATTAA
```

This corresponds to the amino acid sequence <SEQ ID 3058; ORF 576>:

```
m576.pep.. (partial)
  1  ..MQQASYAMGV DIGRSLKQMK EQGAEIDLKV FTEAMQAVYD GKEIKMTEEQ
 51    AQEVMMKFLQ EQQAKAVEKH KADAKANKEK GEAFLKENAA KDGVKTTASG
101    LQYKITKQGE GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPVTFPLSQ
151    VIPGWTEGVQ LLKEGGEATF YIPSNLAYRE QGAGDKIGPN ATLVFDVKLV
201    KIGAPENAPA KQPAQVDIKK VN*
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3059>:

```
g576.seq..(partial)
  1  ..atgggcgtgg acatcggacg ctccctgaaa caaatgaagg aacagggcgc
 51    ggaaatcgat ttgaaagtct ttaccgatgc catgcaggca gtgtatgacg
101    gcaaagaaat caaaatgacc gaagagcagg cccaggaagt gatgatgaaa
151    ttcctgcagg agcagcaggc taaagccgta gaaaaacaca aggcggatgc
201    gaaggccaac aaagaaaaag gcgaagcctt cctgaaggaa aatgccgccg
251    aagacggcgt gaagaccact gcttccggtc tgcagtacaa aatcaccaaa
301    cagggtgaag gcaaacagcc gacaaaagac gacatcgtta ccgtggaata
351    cgaaggccgc ctgattgacg gtaccgtatt cgacagcagc aaagccaacg
401    gcggcccggc caccttccct ttgagccaag tgattccggg ttggaccgaa
451    ggcgtacggc ttctgaaaga aggcggcgaa gccacgttct acatcccgtc
501    caaccttgcc taccgcgaac agggtgcggg cgaaaaaatc ggtccgaacg
551    ccactttggt atttgacgtg aaactggtca aaatcggcgc acccgaaaac
601    gcgcccgcca agcagccgga tcaagtcgac atcaaaaaag taaattaa
```

This corresponds to the amino acid sequence <SEQ ID 3060; ORF 576.ng>:

```
g576.pep..(partial)
  1  ..MGVDIGRSLK QMKEQGAEID LKVFTDAMQA VYDGKEIKMT EEQAQEVMMK
 51    FLQEQQAKAV EKHKADAKAN KEKGEAFLKE NAAEDGVKTT ASGLQYKITK
101    QGEGKQPTKD DIVTVEYEGR LIDGTVFDSS KANGGPATFP LSQVIPGWTE
151    GVRLLKEGGE ATFYIPSNLA YREQGAGEKI GPNATLVFDV KLVKIGAPEN
201    APAKQPDQVD IKKVN*
```

Computer analysis of this amino acid sequence gave the following results:
<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

```
m576/g576  ORFs 576 and 576.ng showed a 97.2% identity in 215 aa overlap

                    10        20        30        40        50        60
m576.pep    MQQASYAMGVDIGRSLKQMKEQGAEIDLKVFTEAMQAVYDGKEIKMTEEQAQEVMMKFLQ
                    |||||||||||||||||||||||:|||||||||||||||||||||||||||
```

```
g576              MGVDIGRSLKQMKEQGAEIDLKVFTDAMQAVYDGKEIKMTEEQAQEVMMKFLQ
                          10        20        30        40        50

                    70        80        90       100       110       120
m576.pep    EQQAKAVEKHKADAKANKEKGEAFLKENAAKDGVKTTASGLQYKITKQGEGKQPTKDDIV
            ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
g576        EQQAKAVEKHKADAKANKEKGEAFLKENAAEDGVKTTASGLQYKITKQGEGKQPTKDDIV
                  60        70        80        90       100       110

                   130       140       150       160       170       180
m576.pep    TVEYEGRLIDGTVFDSSKANGGPVTFPLSQVIPGWTEGVQLLKEGGEATFYIPSNLAYRE
            |||||||||||||||||||||||:|||||||||||||||:||||||||||||||||||||
g576        TVEYEGRLIDGTVFDSSKANGGPATFPLSQVIPGWTEGVRLLKEGGEATFYIPSNLAYRE
                 120       130       140       150       160       170

                   190       200       210       220
m576.pep    QGAGDKIGPNATLVFDVKLVKIGAPENAPAKQPAQVDIKKVNX
            ||||:|||||||||||||||||||||||||||| |||||||||
g576        QGAGEKIGPNATLVFDVKLVKIGAPENAPAKQPDQVDIKKVNX
                 180       190       200       210
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3061>:

```
a576.seq
      1  ATGAACACCA TTTTCAAAAT CAGCGCACTG ACCCTTTCCG CCGCTTTGGC
     51  ACTTTCCGCC TGCGGCAAAA AAGAAGCCGC CCCCGCATCT GCATCCGAAC
    101  CTGCCGCCGC TTCTTCCGCG CAGGGCGACA CCTCTTCGAT CGGCAGCACG
    151  ATGCAGCAGG CAAGCTATGC GATGGGCGTG GACATCGGAC GCTCCCTGAA
    201  GCAAATGAAG GAACAGGGCG CGGAAATCGA TTTGAAAGTC TTTACCGAAG
    251  CCATGCAGGC AGTGTATGAC GGCAAAGAAA TCAAAATGAC CGAAGAGCAG
    301  GCTCAGGAAG TCATGATGAA ATTCCTTCAG GAACAACAGG CTAAAGCCGT
    351  AGAAAAACAC AAGGCGGACG CGAAGGCCAA TAAAGAAAAA GGCGAAGCCT
    401  TTCTGAAAGA AAATGCCGCC AAAGACGGCG TGAAGACCAC TGCTTCCGGC
    451  CTGCAATACA AAATCACCAA ACAGGGCGAA GGCAAACAGC CGACCAAAGA
    501  CGACATCGTT ACCGTGGAAT ACGAAGGCCG CCTGATTGAC GGTACGGTAT
    551  TCGACAGCAG CAAAGCCAAC GGCGGCCCGG TCACCTTCCC TTTGAGCCAA
    601  GTGATTCTGG GTTGGACCGA AGGCGTACAG CTTCTGAAAG AAGGCGGCGA
    651  AGCCACGTTC TACATCCCGT CCAACCTTGC CTACCGCGAA CAGGGTGCGG
    701  GCGACAAAAT CGGCCCGAAC GCCACTTTGG TATTTGATGT GAAACTGGTC
    751  AAAATCGGCG CACCCGAAAA CGCGCCCGCC AAGCAGCCGG CTCAAGTCGA
    801  CATCAAAAAA GTAAATTAA
```

This corresponds to the amino acid sequence <SEQ ID 3062; ORF 576.a>:

```
a576.pep
      1  MNTIFKISAL TLSAALALSA CGKKEAAPAS ASEPAAASSA QGDTSSIGST
     51  MQQASYAMGV DIGRSLKQMK EQGAEIDLKV FTEAMQAVYD GKEIKMTEEQ
    101  AQEVMMKFLQ EQQAKAVEKH KADAKANKEK GEAFLKENAA KDGVKTTASG
    151  LQYKITKQGE GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPVTFPLSQ
    201  VILGWTEGVQ LLKEGGEATF YIPSNLAYRE QGAGDKIGPN ATLVFDVKLV
    251  KIGAPENAPA KQPAQVDIKK VN*
```

```
m576/a576     ORFs 576 and 576.a showed a 99.5% identity in 222 aa overlap

                                                  10        20        30
m576.pep                                MQQASYAMGVDIGRSLKQMKEQGAEIDLKV
                                        ||||||||||||||||||||||||||||||
a576          CGKKEAAPASASEPAAASSAQGDTSSIGSTMQQASYAMGVDIGRSLKQMKEQGAEIDLKV
                      30        40        50        60        70        80

                      40        50        60        70        80        90
m576.pep      FTEAMQAVYDGKEIKMTEEQAQEVMMKFLQEQQAKAVEKHKADAKANKEKGEAFLKENAA
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a576          FTEAMQAVYDGKEIKMTEEQAQEVMMKFLQEQQAKAVEKHKADAKANKEKGEAFLKENAA
                      90       100       110       120       130       140
```

```
                    100       110       120       130       140       150
m576.pep     KDGVKTTASGLQYKITKQGEGKQPTKDDIVTVEYEGRLIDGTVFDSSKANGGPVTFPLSQ
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a576         KDGVKTTASGLQYKITKQGEGKQPTKDDIVTVEYEGRLIDGTVFDSSKANGGPVTFPLSQ
                    150       160       170       180       190       200

                    160       170       180       190       200       210
m576.pep     VIPGWTEGVQLLKEGGEATFYIPSNLAYREQGAGDKIGPNATLVFDVKLVKIGAPENAPA
             || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a576         VILGWTEGVQLLKEGGEATFYIPSNLAYREQGAGDKIGPNATLVFDVKLVKIGAPENAPA
                    210       220       230       240       250       260

                    220
m576.pep     KQPAQVDIKKVNX
             |||||||||||||
a576         KQPAQVDIKKVNX
                    270
```

Further work revealed the DNA sequence identified in *N. meningitidis* <SEQ ID 3063>:

```
m576-1.seq
       1  ATGAACACCA TTTTCAAAAT CAGCGCACTG ACCCTTTCCG CCGCTTTGGC
      51  ACTTTCCGCC TGCGGCAAAA AAGAAGCCGC CCCCGCATCT GCATCCGAAC
     101  CTGCCGCCGC TTCTTCCGCG CAGGGCGACA CCTCTTCGAT CGGCAGCACG
     151  ATGCAGCAGG CAAGCTATGC GATGGGCGTG GACATCGGAC GCTCCCTGAA
     201  GCAAATGAAG GAACAGGGCG CGGAAATCGA TTTGAAAGTC TTTACCGAAG
     251  CCATGCAGGC AGTGTATGAC GGCAAAGAAA TCAAAATGAC CGAAGAGCAG
     301  GCTCAGGAAG TCATGATGAA ATTCCTTCAG GAACAACAGG CTAAAGCCGT
     351  AGAAAAACAC AAGGCGGACG CGAAGGCCAA TAAAGAAAAA GGCGAAGCCT
     401  TTCTGAAAGA AAATGCCGCC AAAGACGGCG TGAAGACCAC TGCTTCCGGC
     451  CTGCAATACA AAATCACCAA ACAGGGCGAA GGCAAACAGC CGACCAAAGA
     501  CGACATCGTT ACCGTGGAAT ACGAAGGCCG CCTGATTGAC GGTACGGTAT
     551  TCGACAGCAG CAAAGCCAAC GGCGGCCCGG TCACCTTCCC TTTGAGCCAA
     601  GTGATTCCGG GTTGGACCGA AGGCGTACAG CTTCTGAAAG AAGGCGGCGA
     651  AGCCACGTTC TACATCCCGT CCAACCTTGC CTACCGCGAA CAGGGTGCGG
     701  GCGACAAAAT CGGTCCGAAC GCCACTTTGG TATTTGATGT GAAACTGGTC
     751  AAAATCGGCG CACCCGAAAA CGCGCCCGCC AAGCAGCCGG CTCAAGTCGA
     801  CATCAAAAAA GTAAATTAA
```

This corresponds to the amino acid sequence <SEQ ID 3064; ORF 576-1>:

```
m576-1.pep
       1  MNTIFKISAL TLSAALALSA CGKKEAAPAS ASEPAAASSA QGDTSSIGST
      51  MQQASYAMGV DIGRSLKQMK EQGAEIDLKV FTEAMQAVYD GKEIKMTEEQ
     101  AQEVMMKFLQ EQQAKAVEKH KADAKANKEK GEAFLKENAA KDGVKTTASG
     151  LQYKITKQGE GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPVTFPLSQ
     201  VIPGWTEGVQ LLKEGGEATF YIPSNLAYRE QGAGDKIGPN ATLVFDVKLV
     251  KIGAPENAPA KQPAQVDIKK VN*
```

The following DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3065>:

```
g576-1.seq
       1  ATGAACACCA TTTTCAAAAT CAGCGCACTG ACCCTTTCCG CCGCTTTGGC
      51  ACTTTCCGCC TGCGGCAAAA AAGAAGCCGC CCCCGCATCT GCATCCGAAC
     101  CTGCCGCCGC TTCTGCCGCG CAGGGCGACA CCTCTTCAAT CGGCAGCACG
     151  ATGCAGCAGG CAAGCTATGC AATGGGCGTG GACATCGGAC GCTCCCTGAA
     201  ACAAATGAAG GAACAGGGCG CGGAAATCGA TTTGAAAGTC TTTACCGATG
     251  CCATGCAGGC AGTGTATGAC GGCAAAGAAA TCAAAATGAC CGAAGAGCAG
     301  GCCCAGGAAG TGATGATGAA ATTCCTGCAG GAGCAGCAGG CTAAAGCCGT
     351  AGAAAAACAC AAGGCGGATG CGAAGGCCAA CAAAGAAAAA GGCGAAGCCT
     401  TCCTGAAGGA AAATGCCGCC AAAGACGGCG TGAAGACCAC TGCTTCCGGT
     451  CTGCAGTACA AAATCACCAA ACAGGGTGAA GGCAAACAGC CGACAAAAGA
     501  CGACATCGTT ACCGTGGAAT ACGAAGGCCG CCTGATTGAC GGTACCGTAT
     551  TCGACAGCAG CAAAGCCAAC GGCGGCCCGG CCACCTTCCC TTTGAGCCAA
     601  GTGATTCCGG GTTGGACCGA AGGCGTACGG CTTCTGAAAG AAGGCGGCGA
     651  AGCCACGTTC TACATCCCGT CCAACCTTGC CTACCGCGAA CAGGGTGCGG
```

```
701  GCGAAAAAAT CGGTCCGAAC GCCACTTTGG TATTTGACGT GAAACTGGTC
751  AAAATCGGCG CACCCGAAAA CGCGCCCGCC AAGCAGCCGG ATCAAGTCGA
801  CATCAAAAAA GTAAATTAA
```

This corresponds to the amino acid sequence <SEQ ID 3066; ORF 576-1.ng>:

```
g576-1.pep
  1  MNTIFKISAL TLSAALALSA CGKKEAAPAS ASEPAAASAA QGDTSSIGST
 51  MQQASYAMGV DIGRSLKQMK EQGAEIDLKV FTDAMQAVYD GKEIKMTEEQ
101  AQEVMMKFLQ EQQAKAVEKH KADAKANKEK GEAFLKENAA KDGVKTTASG
151  LQYKITKQGE GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPATFPLSQ
201  VIPGWTEGVR LLKEGGEATF YIPSNLAYRE QGAGEKIGPN ATLVFDVKLV
251  KIGAPENAPA KQPDQVDIKK VN*
```

g576-1/m576-1 ORFs 576-1 and 576-1.ng showed a 97.8% identity in 272 aa overlap

```
                    10        20        30        40        50        60
g576-1.pep   MNTIFKISALTLSAALALSACGKKEAAPASASEPAAASAAQGDTSSIGSTMQQASYAMGV
             |||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
m576-1       MNTIFKISALTLSAALALSACGKKEAAPASASEPAAASSAQGDTSSIGSTMQQASYAMGV
                    10        20        30        40        50        60

                    70        80        90       100       110       120
g576-1.pep   DIGRSLKQMKEQGAEIDLKVFTDAMQAVYDGKEIKMTEEQAQEVMMKFLQEQQAKAVEKH
             |||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
m576-1       DIGRSLKQMKEQGAEIDLKVFTEAMQAVYDGKEIKMTEEQAQEVMMKFLQEQQAKAVEKH
                    70        80        90       100       110       120

                   130       140       150       160       170       180
g576-1.pep   KADAKANKEKGEAFLKENAAKDGVKTTASGLQYKITKQGEGKQPTKDDIVTVEYEGRLID
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m576-1       KADAKANKEKGEAFLKENAAKDGVKTTASGLQYKITKQGEGKQPTKDDIVTVEYEGRLID
                   130       140       150       160       170       180

                   190       200       210       220       230       240
g576-1.pep   GTVFDSSKANGGPATFPLSQVIPGWTEGVRLLKEGGEATFYIPSNLAYREQGAGEKIGPN
             ||||||||||||:|||||||||||||||:||||||||||||||||||||||||||:||||
m576-1       GTVFDSSKANGGPVTFPLSQVIPGWTEGVQLLKEGGEATFYIPSNLAYREQGAGDKIGPN
                   190       200       210       220       230       240

                   250       260       270
g576-1.pep   ATLVFDVKLVKIGAPENAPAKQPDQVDIKKVNX
             ||||||||||||||||||||||:| ||||||||
m576-1       ATLVFDVKLVKIGAPENAPAKQPAQVDIKKVNX
                   250       260       270
```

The following DNA sequence was identified in *N. meningitidis* <SEQ ID 3067>:

```
a576-1.seq
  1  ATGAACACCA TTTTCAAAAT CAGCGCACTG ACCCTTTCCG CCGCTTTGGC
 51  ACTTTCCGCC TGCGGCAAAA AAGAAGCCGC CCCCGCATCT GCATCCGAAC
101  CTGCCGCCGC TTCTTCCGCG CAGGGCGACA CCTCTTCGAT CGGCAGCACG
151  ATGCAGCAGG CAAGCTATGC GATGGGCGTG GACATCGGAC GCTCCCTGAA
201  GCAAATGAAG GAACAGGGCG CGGAAATCGA TTTGAAAGTC TTTACCGAAG
251  CCATGCAGGC AGTGTATGAC GGCAAAGAAA TCAAAATGAC CGAAGAGCAG
301  GCTCAGGAAG TCATGATGAA ATTCCTTCAG GAACAACAGG CTAAAGCCGT
351  AGAAAAACAC AAGGCGGACG CGAAGGCCAA TAAAGAAAAA GGCGAAGCCT
401  TTCTGAAAGA AAATGCCGCC AAAGACGGCG TGAAGACCAC TGCTTCCGGC
451  CTGCAATACA AAATCACCAA ACAGGGCGAA GGCAAACAGC CGACCAAAGA
501  CGACATCGTT ACCGTGGAAT ACGAAGGCCG CCTGATTGAC GGTACGGTAT
551  TCGACAGCAG CAAAGCCAAC GGCGGCCCGG TCACCTTCCC TTTGAGCCAA
601  GTGATTCTGG GTTGGACCGA AGGCGTACAG CTTCTGAAAG AAGGCGGCGA
651  AGCCACGTTC TACATCCCGT CCAACCTTGC CTACCGCGAA CAGGGTGCGG
```

```
701  GCGACAAAAT CGGCCCGAAC GCCACTTTGG TATTTGATGT GAAACTGGTC
751  AAAATCGGCG CACCCGAAAA CGCGCCCGCC AAGCAGCCGG CTCAAGTCGA
801  CATCAAAAAA GTAAATTAA
```

This corresponds to the amino acid sequence <SEQ ID 3068; ORF 576-1.a>:

```
a576-1.pep
  1  MNTIFKISAL TLSAALALSA CGKKEAAPAS ASEPAAASSA QGDTSSIGST
 51  MQQASYAMGV DIGRSLKQMK EQGAEIDLKV FTEAMQAVYD GKEIKMTEEQ
101  AQEVMMKFLQ EQQAKAVEKH KADAKANKEK GEAFLKENAA KDGVKTTASG
151  LQYKITKQGE GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPVTFPLSQ
201  VILGWTEGVQ LLKEGGEATF YIPSNLAYRE QGAGDKIGPN ATLVFDVKLV
251  KIGAPENAPA KQPAQVDIKK VN*
```

a576-1/m576-1 ORFs 576-1 and 576-1.a showed a 99.6% identity in 272 aa overlap

```
                    10        20        30        40        50        60
a576-1.pep  MNTIFKISALTLSAALALSACGKKEAAPASASEPAAASSAQGDTSSIGSTMQQASYAMGV
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m576-1      MNTIFKISALTLSAALALSACGKKEAAPASASEPAAASSAQGDTSSIGSTMQQASYAMGV
                    10        20        30        40        50        60

                    70        80        90       100       110       120
a576-1.pep  DIGRSLKQMKEQGAEIDLKVFTEAMQAVYDGKEIKMTEEQAQEVMMKFLQEQQAKAVEKH
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m576-1      DIGRSLKQMKEQGAEIDLKVFTEAMQAVYDGKEIKMTEEQAQEVMMKFLQEQQAKAVEKH
                    70        80        90       100       110       120

                   130       140       150       160       170       180
a576-1.pep  KADAKANKEKGEAFLKENAAKDGVKTTASGLQYKITKQGEGKQPTKDDIVTVEYEGRLID
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m576-1      KADAKANKEKGEAFLKENAAKDGVKTTASGLQYKITKQGEGKQPTKDDIVTVEYEGRLID
                   130       140       150       160       170       180

                   190       200       210       220       230       240
a576-1.pep  GTVFDSSKANGGPVTFPLSQVILGWTEGVQLLKEGGEATFYIPSNLAYREQGAGDKIGPN
            |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
m576-1      GTVFDSSKANGGPVTFPLSQVIPGWTEGVQLLKEGGEATFYIPSNLAYREQGAGDKIGPN
                   190       200       210       220       230       240

                   250       260       270
a576-1.pep  ATLVFDVKLVKIGAPENAPAKQPAQVDIKKVNX
            |||||||||||||||||||||||||||||||||
m576-1      ATLVFDVKLVKIGAPENAPAKQPAQVDIKKVNX
                   250       260       270
```

919 gnm43.seq

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 3069>:

```
m919.seq
  1  ATGAAAAAAT ACCTATTCCG CGCCGCCCTG TACGGCATCG CCGCCGCCAT
 51  CCTCGCCGCC TGCCAAAGCA AGAGCATCCA AACCTTTCCG CAACCCGACA
101  CATCCGTCAT CAACGGCCCG GACCGGCCGG TCGGCATCCC CGACCCCGCC
151  GGAACGACGG TCGGCGGCGG CGGGGCCGTC TATACCGTTG TACCGCACCT
201  GTCCCTGCCC CACTGGGCGG CGCAGGATTT CGCCAAAAGC CTGCAATCCT
251  TCCGCCTCGG CTGCGCCAAT TTGAAAAACC GCCAAGGCTG GCAGGATGTG
301  TGCGCCCAAG CCTTTCAAAC CCCCGTCCAT TCCTTTCAGG CAAAACAGTT
351  TTTTGAACGC TATTTCACGC CGTGGCAGGT TGCAGGCAAC GGAAGCCTTG
401  CCGGTACGGT TACCGGCTAT TACGAACCGG TGCTGAAGGG CGACGACAGG
```

```
 451  CGGACGGCAC AAGCCCGCTT CCCGATTTAC GGTATTCCCG ACGATTTTAT
 501  CTCCGTCCCC CTGCCTGCCG GTTTGCGGAG CGGAAAAGCC CTTGTCCGCA
 551  TCAGGCAGAC GGGAAAAAAC AGCGGCACAA TCGACAATAC CGGCGGCACA
 601  CATACCGCCG ACCTCTCCcG ATTCCCCATC ACCGCGCGCA CAACAGCAAT
 651  CAAAGGCAGG TTTGAAGGAA GCCGCTTCCT CCCCTACCAC ACGCGCAACC
 701  AAATCAACGG CGGCGCGCTT GACGGCAAAG CCCCGATACT CGGTTACGCC
 751  GAAGACCCTG TCGAACTTTT TTTTATGCAC ATCCAAGGCT CGGGCCGTCT
 801  GAAAACCCCG TCCGGCAAAT ACATCCGCAT CGGCTATGCC GACAAAAACG
 851  AACATCCyTA CGTTTCCATC GGACGCTATA TGGCGGATAA GGGCTACCTC
 901  AAACTCGGAC AAACCTCCAT GCAGGGCATT AAGTCTTATA TGCGGCAAAA
 951  TCCGCAACGC CTCGCCGAAG TTTTGGGTCA AAACCCCAGC TATATCTTTT
1001  TCCGCGAGCT TGCCGGAAGC AGCAATGACG GCCCTGTCGG CGCACTGGGC
1051  ACGCCGCTGA TGGGGGAATA TGCCGGCGCA GTCGACCGGC ACTACATTAC
1101  CTTGGGTGCG CCCTTATTTG TCGCCACCGC CCATCCGGTT ACCCGCAAAG
1151  CCCTCAACCG CCTGATTATG GCGCAGGATA CCGGCAGCGC GATTAAAGGC
1201  GCGGTGCGCG TGGATTATTT TTGGGGATAC GGCGACGAAG CCGGCGAACT
1251  TGCCGGCAAA CAGAAAACCA CGGGATATGT CTGGCAGCTC CTACCCAACG
1301  GTATGAAGCC CGAATACCGc CCGTAA
```

This corresponds to the amino acid sequence <SEQ ID 3070; ORF 919>:

```
m919.pep
   1  MKKYLFRAAL YGIAAAILAA CQSKSIQTFP QPDTSVINGP DRPVGIPDPA
  51  GTTVGGGGAV YTVVPHLSLP HWAAQDFAKS LQSFRLGCAN LKNRQGWQDV
 101  CAQAFQTPVH SFQAKQFFER YFTPWQVAGN GSLAGTVTGY YEPVLKGDDR
 151  RTAQARFPIY GIPDDFISVP LPAGLRSGKA LVRIRQTGKN SGTIDNTGGT
 201  HTADLSRFPI TARTTAIKGR FEGSRFLPYH TRNQINGGAL DGKAPILGYA
 251  EDPVELFFMH IQGSGRLKTP SGKYIRIGYA DKNEHPYVSI GRYMADKGYL
 301  KLGQTSMQGI KSYMRQNPQR LAEVLGQNPS YIFFRELAGS SNDGPVGALG
 351  TPLMGEYAGA VDRHYITLGA PLFVATAHPV TRKALNRLIM AQDTGSAIKG
 401  AVRVDYFWGY GDEAGELAGK QKTTGYVWQL LPNGMKPEYR P*
```

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 3071>:

```
g919.seq
   1  ATGAAAAAAC ACCTGCTCCG CTCCGCCCTG TACGGcatCG CCGCCgccAT
  51  CctcgCCGCC TGCCAAAgca gGAGCATCCA AACCTTTCCG CAACCCGACA
 101  CATCCGTCAT CAACGGCCCG GACCGGCCGG CCGGCATCCC CGACCCCGCC
 151  GGAACGACGG TTGCCGGCGG CGGGGCCGTC TATACCGTTG TGCCGCACCT
 201  GTCCATGCCC CACTGGGCGG CGCaggATTT TGCCAAAAGC CTGCAATCCT
 251  TCCGCCTCGG CTGCGCCAAT TTGAAAAACC GCCAAGGCTG GCAGGATGTG
 301  TGCGCCCAAG CCTTTCAAAC CCCCGTGCAT TCCTTTCAGG CAAAGcGgTT
 351  TTTTGAACGC TATTTCACGC cgtGGCaggt tgcaggcaAC GGAAGcCTTG
 401  Caggtacggt TACCGGCTAT TACGAACCGG TGCTGAAGGG CGACGGCAGG
 451  CGGACGGAAC GGGCCCGCTT CCCGATTTAC GGTATTCCCG ACGATTTTAT
 501  CTCCGTCCCG CTGCCTGCCG GTTTGCGGGG CGGAAAAAAC CTTGTCCGCA
 551  TCAGGCAGac ggGGAAAAAC AGCGGCACGA TCGACAATGC CGGCGGCACG
 601  CATACCGCCG ACCTCTCCCG ATTCCCCATC ACCGCGCGCA CAACGGcaat
 651  caaaGGCAGG TTTGAaggAA GCCGCTTCCT CCCTTACCAC ACGCGCAACC
 701  AAAtcaacGG CGGCgcgcTT GACGGCAAag ccccCATCCT CggttacgcC
 751  GAagaccCcG tcgaacttTT TTTCATGCAC AtccaaggCT CGGGCCGCCT
 801  GAAAACCCcg tccggcaaat acatCCGCAt cggaTacgcc gacAAAAACG
 851  AACAtccgTa tgtttccatc ggACGctaTA TGGCGGACAA AGGCTACCTC
 901  AAGctcgggc agACCTCGAT GCAGGgcatc aaagcCTATA TGCGGCAAAA
 951  TCCGCAACGC CTCGCCGAAG TTTTGGGTCA AAACCCCAGC TATATCTTTT
1001  TCCGCGAGCT TGCCGGAAGC GGCAATGAGG GCCCCGTCGG CGCACTGGGC
1051  ACGCCACTGA TGGGGGAATA CGCCGGCGCA ATCGACCGGC ACTACATTAC
1101  CTTGGGCGCG CCCTTATTTG TCGCCACCGC CCATCCGGTT ACCCGCAAAG
1151  CCCTCAACCG CCTGATTATG GCGCAGGATA CAGGCAGCGC GATCAAAGGC
1201  GCGGTGCGCG TGGATTATTT TTGGGGTTAC GGCGACGAAG CCGGCGAACT
1251  TGCCGGCAAA CAGAAAACCA CGGGATACGT CTGGCAGCTC CTGCCCAACG
1301  GCATGAAGCC CGAATACCGC CCGTGA
```

This corresponds to the amino acid sequence <SEQ ID 3072; ORF 919.ng>:

```
g919.pep
  1  MKKHLLRSAL YGIAAAILAA CQSRSIQTFP QPDTSVINGP DRPAGIPDPA
 51  GTTVAGGGAV YTVVPHLSMP HWAAQDFAKS LQSFRLGCAN LKNRQGWQDV
101  CAQAFQTPVH SFQAKRFFER YFTPWQVAGN GSLAGTVTGY YEPVLKGDGR
151  RTERARFPIY GIPDDFISVP LPAGLRGGKN LVRIRQTGKN SGTIDNAGGT
201  HTADLSRFPI TARTTAIKGR FEGSRFLPYH TRNQINGGAL DGKAPILGYA
251  EDPVELFFMH IQGSGRLKTP SGKYIRIGYA DKNEHPYVSI GRYMADKGYL
301  KLGQTSMQGI KAYMRQNPQR LAEVLGQNPS YIFFRELAGS GNEGPVGALG
351  TPLMGEYAGA IDRHYITLGA PLFVATAHPV TRKALNRLIM AQDTGSAIKG
401  AVRVDYFWGY GDEAGELAGK QKTTGYVWQL LPNGMKPEYR P*
```

ORF 919 shows 95.9 % identity over a 441 aa overlap with a predicted ORF (ORF 919.ng) from *N. gonorrhoeae:*

```
m919/g919

                   10        20        30        40        50        60
m919.pep   MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
           |||:|:|:|||||||||||||||:||||||||||||||||||:||||||||||||:|||||
g919       MKKHLLRSALYGIAAAILAACQSRSIQTFPQPDTSVINGPDRPAGIPDPAGTTVAGGGAV
                   10        20        30        40        50        60

                   70        80        90       100       110       120
m919.pep   YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
           ||||||||:||||||||||||||||||||||||||||||||||||||||||||||:||||
g919       YTVVPHLSMPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKRFFER
                   70        80        90       100       110       120

                  130       140       150       160       170       180
m919.pep   YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
           |||||||||||||||||||||||||||| |||  :||||||||||||||||||||||:||
g919       YFTPWQVAGNGSLAGTVTGYYEPVLKGDGRRTERARFPIYGIPDDFISVPLPAGLRGGKN
                  130       140       150       160       170       180

                  190       200       210       220       230       240
m919.pep   LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
           ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
g919       LVRIRQTGKNSGTIDNAGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
                  190       200       210       220       230       240

                  250       260       270       280       290       300
m919.pep   DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g919       DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
                  250       260       270       280       290       300

                  310       320       330       340       350       360
m919.pep   KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
           |||||||||||:||||||||||||||||||||||||||||:|:|||||||||||||||||
g919       KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSGNEGPVGALGTPLMGEYAGA
                  310       320       330       340       350       360

                  370       380       390       400       410       420
m919.pep   VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
           :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g919       IDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
                  370       380       390       400       410       420
```

```
                 430       440
m919.pep   QKTTGYVWQLLPNGMKPEYRPX
           ||||||||||||||||||||||
g919       QKTTGYVWQLLPNGMKPEYRPX
                 430       440
```

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 3073>:

```
a919.seq
   1  ATGAAAAAAT ACCTATTCCG CGCCGCCCTG TGCGGCATCG CCGCCGCCAT
  51  CCTCGCCGCC TGCCAAAGCA AGAGCATCCA AACCTTTCCG CAACCCGACA
 101  CATCCGTCAT CAACGGCCCG GACCGGCCGG TCGGCATCCC CGACCCCGCC
 151  GGAACGACGG TCGGCGGCGG CGGGGCCGTT TATACCGTTG TGCCGCACCT
 201  GTCCCTGCCC CACTGGGCGG CGCAGGATTT CGCCAAAAGC CTGCAATCCT
 251  TCCGCCTCGG CTGCGCCAAT TTGAAAAACC GCCAAGGCTG GCAGGATGTG
 301  TGCGCCCAAG CCTTTCAAAC CCCCGTCCAT TCCGTTCAGG CAAAACAGTT
 351  TTTTGAACGC TATTTCACGC CGTGGCAGGT TGCAGGCAAC GGAAGCCTTG
 401  CCGGTACGGT TACCGGCTAT TACGAGCCGG  GCTGAAGGG CGACGACAGG
 451  CGGACGGCAC AAGCCCGCTT CCCGATTTAC GGTATTCCCG ACGATTTTAT
 501  CTCCGTCCCC CTGCCTGCCG GTTTGCGGAG CGGAAAAGCC CTTGTCCGCA
 551  TCAGGCAGAC GGGAAAAAAC AGCGGCACAA TCGACAATAC CGGCGGCACA
 601  CATACCGCCG ACCTCTCCCA ATTCCCCATC ACTGCGCGCA CAACGGCAAT
 651  CAAAGGCAGG TTTGAAGGAA GCCGCTTCCT CCCCTACCAC ACGCGCAACC
 701  AAATCAACGG CGGCGCGCTT GACGGCAAAG CCCCGATACT CGGTTACGCC
 751  GAAGACCCCG TCGAACTTTT TTTTATGCAC ATCCAAGGCT CGGGCCGTCT
 801  GAAAACCCCG TCCGGCAAAT ACATCCGCAT CGGCTATGCC GACAAAAACG
 851  AACATCCCTA CGTTTCCATC GGACGCTATA TGGCGGACAA AGGCTACCTC
 901  AAGCTCGGGC AGACCTCGAT GCAGGGCATC AAAGCCTATA TGCAGCAAAA
 951  CCCGCAACGC CTCGCCGAAG TTTTGGGGCA AAACCCCAGC TATATCTTTT
1001  TCCGAGAGCT TACCGGAAGC AGCAATGACG GCCCTGTCGG CGCACTGGGC
1051  ACGCCGCTGA TGGGCGAGTA CGCCGGCGCA GTCGACCGGC ACTACATTAC
1101  CTTGGGCGCG CCCTTATTTG TCGCCACCGC CCATCCGGTT ACCCGCAAAG
1151  CCCTCAACCG CCTGATTATG GCGCAGGATA CCGGCAGCGC GATTAAAGGC
1201  GCGGTGCGCG TGGATTATTT TTGGGGATAC GGCGACGAAG CCGGCGAACT
1251  TGCCGGCAAA CAGAAAACCA CGGGATATGT CTGGCAGCTT CTGCCCAACG
1301  GTATGAAGCC CGAATACCGC CCGTAA
```

This corresponds to the amino acid sequence <SEQ ID 3074; ORF 919.a>:

```
a919.pep
   1  MKKYLFRAAL CGIAAAILAA CQSKSIQTFP QPDTSVINGP DRPVGIPDPA
  51  GTTVGGGGAV YTVVPHLSLP HWAAQDFAKS LQSFRLGCAN LKNRQGWQDV
 101  CAQAFQTPVH SVQAKQFFER YFTPWQVAGN GSLAGTVTGY YEPVLKGDDR
 151  RTAQARFPIY GIPDDFISVP LPAGLRSGKA LVRIRQTGKN SGTIDNTGGT
 201  HTADLSQFPI TARTTAIKGR FEGSRFLPYH TRNQINGGAL DGKAPILGYA
 251  EDPVELFFMH IQGSGRLKTP SGKYIRIGYA DKNEHPYVSI GRYMADKGYL
 301  KLGQTSMQGI KAYMQQNPQR LAEVLGQNPS YIFFRELTGS SNDGPVGALG
 351  TPLMGEYAGA VDRHYITLGA PLFVATAHPV TRKALNRLIM AQDTGSAIKG
 401  AVRVDYFWGY GDEAGELAGK QKTTGYVWQL LPNGMKPEYR P*
```

m919/a919 ORFs 919 and 919.a showed a 98.6% identity in 441 aa overlap

```
                   10        20        30        40        50        60
m919.pep   MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
           |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
a919       MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
                   10        20        30        40        50        60

                   70        80        90       100       110       120
m919.pep   YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
           |||||||||||||||||||||||||||||||||||||||||:||||||||| ||||||||
a919       YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSVQAKQFFER
                   70        80        90       100       110       120
```

```
                   130       140       150       160       170       180
m919.pep   YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a919       YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m919.pep   LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
           ||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||
a919       LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
                   190       200       210       220       230       240

                   250       260       270       280       290       300
m919.pep   DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a919       DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
                   250       260       270       280       290       300

                   310       320       330       340       350       360
m919.pep   KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
           |||||||||||:||:||||||||||||||||||||||:||||||||||||||||||||||
a919       KLGQTSMQGIKAYMQQNPQRLAEVLGQNPSYIFFRELTGSSNDGPVGALGTPLMGEYAGA
                   310       320       330       340       350       360

                   370       380       390       400       410       420
m919.pep   VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a919       VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
                   370       380       390       400       410       420

                   430       440
m919.pep   QKTTGYVWQLLPNGMKPEYRPX
           |||||||||||||||||||||
a919       QKTTGYVWQLLPNGMKPEYRPX
                   430       440
```

121 and 121-1

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3075>:

m121.seq

```
   1 ATGGAAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
  51 GGCGGATGCC GTACTGATAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
 101 AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTTACGCCG CCAATTGCTG
 151 GATTTGCAGG ACACAGGCGC AGACGAACTG CACCGCAGCA GGATTTTGTC
 201 GCAAGAACTC AGCCGCCTAT ATGCGCAAAC CGCCGCCGAA CTGCTGTGCA
 251 GTCAAAACCT CGCACCGTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
 301 ACCGTCCGAC ACGCGCCGGA ACACGGTTAC AGCATACAGC TTGCCGATTT
 351 GCCGCTGCTG GCGxxxxxxx xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx
 401 xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx
 451 xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx
 501 xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx
 551 xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx
 601 xxxxxxCAGC TTCCTTACGA CAAAAACGGT GCAAAGTCGG CACAAGGCAA
 651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCCACCCG TATTTCGCAC
 701 AACGCCACCC TAAAAGCACG GGGCGCGAAC TGTTTGCCAT AAATTGGCTC
 751 GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TGCGGACGCT
 801 TTCCCGTTTT ACCGCGCAAA CCGTTTGCGA CGCCGTCTCA CACGCAGCGG
 851 CAGATGCCCG TCAAATGTAC ATTTGCGACG GCGGCATCCG CAATCCTGTT
 901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
 951 CACCGCCGAC CTGAACCTCG ATCCGCAATG GGTGGAAGCC GCCGnATTTG
1001 CGTGGTTGGC GGCGTGTTGG ATTAATCGCA TTCCCGGTAG TCCGCACAAA
```

```
1051  GCAACCGGCG CATCCAAACC GTGTATTCTG AnCGCGGGAT ATTATTATTG
1101  A
```

This corresponds to the amino acid sequence <SEQ ID 3076; ORF 121>:

m121.pep

```
  1  METQLYIGIM SGTSMDGADA VLIRMDGGKW LGAEGHAFTP YPGRLRRQLL
 51  DLQDTGADEL HRSRILSQEL SRLYAQTAAE LLCSQNLAPS DITALGCHGQ
101  TVRHAPEHGY SIQLADLPLL Axxxxxxxxx xxxxxxxxxx xxxxxxxxxx
151  xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx
201  xxQLPYDKNG AKSAQGNILP QLLDRLLAHP YFAQRHPKST GRELFAINWL
251  ETYLDGGENR YDVLRTLSRF TAQTVCDAVS HAAADARQMY ICDGGIRNPV
301  LMADLAECFG TRVSLHSTAD LNLDPQWVEA AXFAWLAACW INRIPGSPHK
351  ATGASKPCIL XAGYYY*
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3077>:

g121.seq

```
   1  ATGGAAACAC AGCTTTACAT CGGCATTATG TCGGGAACCA GTATGGACGG
  51  GGCGGATGCC GTGCTGGTAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
 101  AAGGGCACGC CTTTACCCCC TACCCTGACC GGTTGCGCCG CAAATTGCTG
 151  GATTTGCAGG ACACAGGCAC AGACGAACTG CACCGCAGCA GGATGTTGTC
 201  GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
 251  GTCAAAACCT CGCTCCGTGC GACATTACCG CCCTCGGCTG CCACGGGCAA
 301  ACCGTCCGAC ACGCGCCGGA ACACGGTtac -G [illegible]TACAGC TTGCCGATTT
 351  GCCGCTGCTG GCGGAACTGa cgcggatttT TACCGTCggc gacttcCGCA
 401  GCCGCGACCT TGCTGCCGGC GGacaAGGTG CGCCGCTCGT CCCCGCCTTT
 451  CACGAAGCCC TGTTCCGCGA TGACAGGGAA ACACGCGTGG TACTGAACAT
 501  CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCGGCGCA CCCGCCTTCG
 551  GCTTCGACAC AGGGCCGGGC AATATGCTGA TGGAcgcgtg gacgcaggca
 601  cacTGGcagc TGCCTTACGA CAAAAacggt gcAAAGgcgg cacAAGGCAA
 651  catatTGCcg cAACTGCTCG gcaggctGCT CGCCcaccCG TATTTCTCAC
 701  AACCCcaccc aaAAAGCACG GGgcGCGaac TgtttgcccT AAattggctc
 751  gaaacctAcc ttgacggcgg cgaaaaccga tacgacgtat tgcggacgct
 801  ttcccgattc accgcgcaaA ccgTttggga cgccgtctca CACGCAGCGG
 851  CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
 901  TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
 951  CACCGCCGAA CTGAACCTCG ATCCTCAATG GGTGGAGGCG gccgCATTtg
1001  cgtggttggC GGCGTGTTGG ATTAACCGCA TTCCCGGTAG TCCGCACAAA
1051  GCGACCGGCG CATCCAAACC GTGTATTCTG GGCGCGGGAT ATTATTATTG
1101  A
```

This corresponds to the amino acid sequence <SEQ ID 3078; ORF 121.ng>:

g121.pep

```
  1  METQLYIGIM SGTSMDGADA VLVRMDGGKW LGAEGHAFTP YPDRLRRKLL
 51  DLQDTGTDEL HRSRMLSQEL SRLYAQTAAE LLCSQNLAPC DITALGCHGQ
101  TVRHAPEHGY SIQLADLPLL AELTRIFTVG DFRSRDLAAG GQGAPLVPAF
151  HEALFRDDRE TRVVLNIGGI ANISVLPPGA PAFGFDTGPG NMLMDAWTQA
201  HWQLPYDKNG AKAAQGNILP QLLGRLLAHP YFSQPHPKST GRELFALNWL
251  ETYLDGGENR YDVLRTLSRF TAQTVWDAVS HAAADARQMY ICGGGIRNPV
301  LMADLAECFG TRVSLHSTAE LNLDPQWVEA AAFAWLAACW INRIPGSPHK
351  ATGASKPCIL GAGYYY*
```

ORF 121 shows 73.5% identity over a 366 aa overlap with a predicted ORF (ORF121.ng) from *N. gonorrhoeae:*

m121/g121

```
                   10        20        30        40        50        60
m121.pep   METQLYIGIMSGTSMDGADAVLIRMDGGKWLGAEGHAFTPYPGRLRRQLLDLQDTGADEL
           ||||||||||||||||||||||:|||||||||||||||||||| ||||:|||||||:|||
g121       METQLYIGIMSGTSMDGADAVLVRMDGGKWLGAEGHAFTPYPDRLRRKLLDLQDTGTDEL
                   10        20        30        40        50        60
                   70        80        90       100       110       120
m121.pep   HRSRILSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPEHGYSIQLADLPLL
           ||||:|||||||||||||||||||||||| ||||||||||||||||||||||||||||||
```

```
g121            HRSRMLSQELSRLYAQTAAELLCSQNLAPCDITALGCHGQTVRHAPEHGYSIQLADLPLL
                        70        80        90       100       110       120
                       130       140       150       160       170       180
m121.pep        AXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
                |  :   :                                :
g121            AELTRIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDDRETRVVLNIGGIANISVLPPGA
                       130       140       150       160       170       180
                       190       200       210       220       230       240
m121.pep        XXXXXXXXXXXXXXXXXXXXXXQLPYDKNGAKSAQGNILPQLLDRLLAHPYFAQRHPKST
                      :          :    ||||||||||:|||||||||| ||||||||:| |||||
g121            PAFGFDTGPGNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLGRLLAHPYFSQPHPKST
                       190       200       210       220       230       240
                       250       260       270       280       290       300
m121.pep        GRELFAINWLETYLDGGENRYDVLRTLSRFTAQTVCDAVSHAAADARQMYICDGGIRNPV
                ||||||:|||||||||||||||||||||||||||| |||||||||||||||| |||||||
g121            GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVWDAVSHAAADARQMYICGGGIRNPV
                       250       260       270       280       290       300
                       310       320       330       340       350       360
m121.pep        LMADLAECFGTRVSLHSTADLNLDPQWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL
                |||||||||||||||||||:||||||||||| ||||||||||||||||||||||||||||
g121            LMADLAECFGTRVSLHSTAELNLDPQWVEAAAFAWLAACWINRIPGSPHKATGASKPCIL
                       310       320       330       340       350       360

m121.pep        XAGYYYX
                 ||||||
g121            GAGYYYX
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3079>:

```
a121.seq
       1  ATGGAAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
      51  GGCGGATGCC GTACTGATAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
     101  AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTTACGCCG CAAATTGCTG
     151  GATTTGCAGG ACACAGGCGC GGACGAACTG CACCGCAGCA GGATGTTGTC
     201  GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
     251  GTCAAAACCT CGCGCCGTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
     301  ACCGTCAGAC ACGCGCCGGA ACACAGTTAC AGCGTACAGC TTGCCGATTT
     351  GCCGCTGCTG GCGGAACGGA CTCAGATTTT TACCGTCGGC GACTTCCGCA
     401  GCCGCGACCT TGCGGCCGGC GGACAAGGCG CGCCGCTCGT CCCCGCCTTT
     451  CACGAAGCCC TGTTCCGCGA CGACAGGGAA ACACGCGCGG TACTGAACAT
     501  CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCGACGCA CCCGCCTTCG
     551  GCTTCGACAC AGGACCGGGC AATATGCTGA TGGACGCGTG GATGCAGGCA
     601  CACTGGCAGC TTCCTTACGA CAAAAACGGT GCAAAGGCGG CACAAGGCAA
     651  CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCCACCCG TATTTCGCAC
     701  AACCCCACCC TAAAAGCACG GGGCGCGAAC TGTTTGCCCT AAATTGGCTC
     751  GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TGCGGACGCT
     801  TTCCCGATTC ACCGCGCAAA CCGTTTTCGA CGCCGTCTCA CACGCAGCGG
     851  CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
     901  TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
     951  CACCGCCGAA CTGAACCTCG ATCCGCAATG GGTAGAAGCC GCCGCGTTCG
    1001  CATGGATGGC GGCGTGTTGG GTCAACCGCA TTCCCGGTAG TCCGCACAAA
    1051  GCAACCGGCG CATCCAAACC GTGTATTCTG GGCGCGGGAT ATTATTATTG
    1101  A
```

This corresponds to the amino acid sequence <SEQ ID 3080; ORF 121.a>:

```
a121.pep
       1  METQLYIGIM SGTSMDGADA VLIRMDGGKW LGAEGHAFTP YPGRLRRKLL
      51  DLQDTGADEL HRSRMLSQEL SRLYAQTAAE LLCSQNLAPS DITALGCHGQ
     101  TVRHAPEHSY SVQLADLPLL AERTQIFTVG DFRSRDLAAG GQGAPLVPAF
     151  HEALFRDDRE TRAVLNIGGI ANISVLPPDA PAFGFDTGPG NMLMDAWMQA
     201  HWQLPYDKNG AKAAQGNILP QLLDRLLAHP YFAQPHPKST GRELFALNWL
     251  ETYLDGGENR YDVLRTLSRF TAQTVFDAVS HAAADARQMY ICGGGIRNPV
```

```
301  LMADLAECFG TRVSLHSTAE LNLDPQWVEA AAFAWMAACW VNRIPGSPHK
351  ATGASKPCIL GAGYYY*
```

m121/a121 ORFs 121 and 121.a showed a 74.0% identity in 366 aa overlap

```
                      10        20        30        40        50        60
m121.pep    METQLYIGIMSGTSMDGADAVLIRMDGGKWLGAEGHAFTPYPGRLRRQLLDLQDTGADEL
            |||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
a121        METQLYIGIMSGTSMDGADAVLIRMDGGKWLGAEGHAFTPYPGRLRRKLLDLQDTGADEL
                      10        20        30        40        50        60

                      70        80        90       100       110       120
m121.pep    HRSRILSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPEHGYSIQLADLPLL
            ||||:|||||||||||||||||||||||||||||||||||||||||||||:||:||||||||
a121        HRSRMLSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPEHSYSVQLADLPLL
                      70        80        90       100       110       120

                     130       140       150       160       170       180
m121.pep    AXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
            |  :    :                                      :
a121        AERTQIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDDRETRAVLNIGGIANISVLPPDA
                     130       140       150       160       170       180

                     190       200       210       220       230       240
m121.pep    XXXXXXXXXXXXXXXXXXXXXXQLPYDKNGAKSAQGNILPQLLDRLLAHPYFAQRHPKST
                :                     ||||||||||:|||||||||||||||||||||| |||||
a121        PAFGFDTGPGNMLMDAWMQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST
                     190       200       210       220       230       240

                     250       260       270       280       290       300
m121.pep    GRELFAINWLETYLDGGENRYDVLRTLSRFTAQTVCDAVSHAAADARQMYICDGGIRNPV
            ||||||:|||||||||||||||||||||||||||||| ||||||||||||||||| |||||||
a121        GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVFDAVSHAAADARQMYICGGGIRNPV
                     250       260       270       280       290       300

                     310       320       330       340       350       360
m121.pep    LMADLAECFGTRVSLHSTADLNLDPQWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL
            ||||||||||||||||||:||||||||||| |||:||||:||||||||||||||||||||
a121        LMADLAECFGTRVSLHSTAELNLDPQWVEAAAFAWMAACWVNRIPGSPHKATGASKPCIL
                     310       320       330       340       350       360

m121.pep    XAGYYYX
             ||||||
a121        GAGYYYX
```

Further work revealed the DNA sequence identified in *N. meningitidis* <SEQ ID 3081>:

```
m121-1.seq
  1  ATGGAAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
 51  GGCGGATGCC GTACTGATAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
101  AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTTACGCCG CCAATTGCTG
151  GATTTGCAGG ACACAGGCGC AGACGAACTG CACCGCAGCA GGATTTTGTC
201  GCAAGAACTC AGCCGCCTAT ATGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251  GTCAAAACCT CGCACCGTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
301  ACCGTCCGAC ACGCGCCGGA ACACGGTTAC AGCATACAGC TTGCCGATTT
351  GCCGCTGCTG GCGGAACGGA CGCGGATTTT TACCGTCGGC GACTTCCGCA
401  GCCGCGACCT TGCGGCCGGC GGACAAGGCG CGCCACTCGT CCCCGCCTTT
451  CACGAAGCCC TGTTCCGCGA CAACAGGGAA ACACGCGCGG TACTGAACAT
501  CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCGACGCA CCCGCCTTCG
551  GCTTCGACAC AGGGCCGGGC AATATGCTGA TGGACGCGTG GACGCAGGCA
601  CACTGGCAGC TTCCTTACGA CAAAAACGGT GCAAAGGCGG CACAAGGCAA
651  CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCCACCCG TATTTCGCAC
701  AACCCCACCC TAAAAGCACG GGGCGCGAAC TGTTTGCCCT AAATTGGCTC
751  GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TGCGGACGCT
801  TTCCCGTTTT ACCGCGCAAA CCGTTTGCGA CGCCGTCTCA CACGCAGCGG
```

```
 851  CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
 901  TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
 951  CACCGCCGAC CTGAACCTCG ATCCGCAATG GGTGGAAGCC GCCGNATTTG
1001  CGTGGTTGGC GGCGTGTTGG ATTAATCGCA TTCCCGGTAG TCCGCACAAA
1051  GCAACCGGCG CATCCAAACC GTGTATTCTG ANCGCGGGAT ATTATTATTG
1101  A
```

This corresponds to the amino acid sequence <SEQ ID 3082; ORF 121-1>:

```
m121-1.pep
  1  METQLYIGIM SGTSMDGADA VLIRMDGGKW LGAEGHAFTP YPGRLRRQLL
 51  DLQDTGADEL HRSRILSQEL SRLYAQTAAE LLCSQNLAPS DITALGCHGQ
101  TVRHAPEHGY SIQLADLPLL AERTRIFTVG DFRSRDLAAG GQGAPLVPAF
151  HEALFRDNRE TRAVLNIGGI ANISVLPPDA PAFGFDTGPG NMLMDAWTQA
201  HWQLPYDKNG AKAAQGNILP QLLDRLLAHP YFAQPHPKST GRELFALNWL
251  ETYLDGGENR YDVLRTLSRF TAQTVCDAVS HAAADARQMY ICGGGIRNPV
301  LMADLAECFG TRVSLHSTAD LNLDPQWVEA AXFAWLAACW INRIPGSPHK
351  ATGASKPCIL XAGYYY*
```

m121-1/g121 ORFs 121-1 and 121.ng showed a 95.6% identity in 366 aa overlap

```
                    10        20        30        40        50        60
m121-1.pep  METQLYIGIMSGTSMDGADAVLIRMDGGKWLGAEGHAFTPYPGRLRRQLLDLQDTGADEL
            ||||||||||||||||||||||:||||||||||||||||||| ||||:||||||||:|||
g121        METQLYIGIMSGTSMDGADAVLVRMDGGKWLGAEGHAFTPYPDRLRRKLLDLQDTGTDEL
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m121-1.pep  HRSRILSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPEHGYSIQLADLPLL
            ||||:||||||||||||||||||||||||| |||||||||||||||||||||||||||||
g121        HRSRMLSQELSRLYAQTAAELLCSQNLAPCDITALGCHGQTVRHAPEHGYSIQLADLPLL
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m121-1.pep  AERTRIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDNRETRAVLNIGGIANISVLPPDA
            || |||||||||||||||||||||||||||||||||||:||||:||||||||||||||| |
g121        AELTRIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDDRETRVVLNIGGIANISVLPPGA
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m121-1.pep  PAFGFDTGPGNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST
            |||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
g121        PAFGFDTGPGNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLGRLLAHPYFSQPHPKST
                   190       200       210       220       230       240

                   250       260       270       280       290       300
m121-1.pep  GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVCDAVSHAAADARQMYICGGGIRNPV
            ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
g121        GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVWDAVSHAAADARQMYICGGGIRNPV
                   250       260       270       280       290       300

                   310       320       330       340       350       360
m121-1.pep  LMADLAECFGTRVSLHSTADLNLDPQWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL
            |||||||||||||||||||:||||||||||| ||||||||||||||||||||||||||||
g121        LMADLAECFGTRVSLHSTAELNLDPQWVEAAAFAWLAACWINRIPGSPHKATGASKPCIL
                   310       320       330       340       350       360

m121-1.pep  XAGYYYX
             ||||||
g121        GAGYYYX
```

The following DNA sequence was identified in *N. meningitidis* <SEQ ID 3083>:

```
a121-1.seq
   1  ATGGAAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
```

```
  51  GGCGGATGCC GTACTGATAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
 101  AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTTACGCCG CAAATTGCTG
 151  GATTTGCAGG ACACAGGCGC GGACGAACTG CACCGCAGCA GGATGTTGTC
 201  GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
 251  GTCAAAACCT CGCGCCGTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
 301  ACCGTCAGAC ACGCGCCGGA ACACAGTTAC AGCGTACAGC TTGCCGATTT
 351  GCCGCTGCTG GCGGAACGGA CTCAGATTTT TACCGTCGGC GACTTCCGCA
 401  GCCGCGACCT TGCGGCCGGC GGACAAGGCG CGCCGCTCGT CCCCGCCTTT
 451  CACGAAGCCC TGTTCCGCGA CGACAGGGAA ACACGCGCGG TACTGAACAT
 501  CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCGACGCA CCCGCCTTCG
 551  GCTTCGACAC AGGACCGGGC AATATGCTGA TGGACGCGTG GATGCAGGCA
 601  CACTGGCAGC TTCCTTACGA CAAAAACGGT GCAAAGGCGG CACAAGGCAA
 651  CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCCACCCG TATTTCGCAC
 701  AACCCCACCC TAAAAGCACG GGGCGCGAAC TGTTTGCCCT AAATTGGCTC
 751  GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TGCGGACGCT
 801  TTCCCGATTC ACCGCGCAAA CCGTTTTCGA CGCCGTCTCA CACGCAGCGG
 851  CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
 901  TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
 951  CACCGCCGAA CTGAACCTCG ATCCGCAATG GGTAGAAGCC GCCGCGTTCG
1001  CATGGATGGC GGCGTGTTGG GTCAACCGCA TTCCCGGTAG TCCGCACAAA
1051  GCAACCGGCG CATCCAAACC GTGTATTCTG GGCGCGGGAT ATTATTATTG
1101  A
```

This corresponds to the amino acid sequence <SEQ ID 3054; ORF 121-1.a>:

```
a121-1.pep
   1  METQLYIGIM SGTSMDGADA VLIRMDGGKW LGAEGHAFTP YPGRLRRKLL
  51  DLQDTGADEL HRSRMLSQEL SRLYAQTAAE LLCSQNLAPS DITALGCHGQ
 101  TVRHAPEHSY SVQLADLPLL AERTQIFTVG DFRSRDLAAG GQGAPLVPAF
 151  HEALFRDDRE TRAVLNIGGI ANISVLPPDA PAFGFDTGPG NMLMDAWMQA
 201  HWQLPYDKNG AKAAQGNILP QLLDRLLAHP YFAQPHPKST GRELFALNWL
 251  ETYLDGGENR YDVLRTLSRF TAQTVFDAVS HAAADARQMY ICGGGIRNPV
 301  LMADLAECFG TRVSLHSTAE LNLDPQWVEA AAFAWMAACW VNRIPGSPHK
 351  ATGASKPCIL GAGYYY*
```

m121-1/a121-1 ORFs 121-1 and 121-1.a showed a 96.4% identity in 366 aa overlap

```
                     10        20        30        40        50        60
m121-1.pep   METQLYIGIMSGTSMDGADAVLIRMDGGKWLGAEGHAFTPYPGRLRRQLLDLQDTGADEL
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
a121-1       METQLYIGIMSGTSMDGADAVLIRMDGGKWLGAEGHAFTPYPGRLRRKLLDLQDTGADEL
                     10        20        30        40        50        60

                     70        80        90       100       110       120
m121-1.pep   HRSRILSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPEHGYSIQLADLPLL
             ||||:|||||||||||||||||||||||||||||||||||||||||||:||:||||||||
a121-1       HRSRMLSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPEHSYSVQLADLPLL
                     70        80        90       100       110       120

                    130       140       150       160       170       180
m121-1.pep   AERTRIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDNRETRAVLNIGGIANISVLPPDA
             ||||:|||||||||||||||||||||||||||||||:|||||||||||||||||||||||
a121-1       AERTQIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDDRETRAVLNIGGIANISVLPPDA
                    130       140       150       160       170       180

                    190       200       210       220       230       240
m121-1.pep   PAFGFDTGPGNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST
             ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
a121-1       PAFGFDTGPGNMLMDAWMQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST
                    190       200       210       220       230       240

                    250       260       270       280       290       300
m121-1.pep   GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVCDAVSHAAADARQMYICGGGIRNPV
             |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
a121-1       GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVFDAVSHAAADARQMYICGGGIRNPV
                    250       260       270       280       290       300
```

```
                  310       320       330       340       350       360
m121-1.pep  LMADLAECFGTRVSLHSTADLNLDPQWVEAAXFAWLAACWINRRIPGSPHKATGASKPCIL
            |||||||||||||||||||:||||||||||| |||:||||:||||||||||||||||||||
a121        LMADLAECFGTRVSLHSTAELNLDPQWVEAAAFAWMAACWVNRRIPGSPHKATGASKPCIL
                  310       320       330       340       350       360

m121-1.pep  XAGYYYX
             ||||||
a121        GAGYYYX
```

128 and 128-1

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3085>:

```
m128.seq  (partial)
     1  ATGACTGACA ACGCACTGCT CCATTTGGGC GAAGAACCCC GTTTTGATCA
    51  AATCAAAACC GAAGACATCA AACCCGCCCT GCAAACCGCC ATCGCCGAAG
   101  CGCGCGAACA AATCGCCGCC ATCAAAGCCC AAACGCACAC CGGCTGGGCA
   151  AACACTGTCG AACCCCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
   201  GGGCGTGGTG TCGCACCTCA ACTGCGTCGC CGACACGCCC GAACTGCGCG
   251  CCGTCTATAA CGAACTGATG CCCGAAATCA CCGTCTTCTT CACCGAAATC
   301  GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
   351  CGAATTCGAC ACCCTCTCCC CCGCACAAAA AACCAAACTC AACCAC
     1  TACGCCAGCG AAAAACTGCG CGAAGCCAAA TACGCGTTCA GCGAAACCGA
    51  wGTCAAAAAA TAyTTCCCyG TCGGCAAwGT ATTAAACGGA CTGTTCGCCC
   101  AAmTCAAAAA ACTmTACGGC ATCGGATTTA CCGAAAAAAC yGTCCCCGTC
   151  TGGCACAAAG ACGTGCGCTA TTkTGAATTG CAACAAAACG GCGAAmCCAT
   201  AGGCGGCGTT TATATGGATT TGTACGCACG CGAAGGCAAA CGCGGCGGCG
   251  CGTGGATGAA CGACTACAAA GGCCGCCGCC GTTTTTCAGA CGGCACGCTG
   301  CAAyTGCCCA CCGCCTACCT CGTCTGCAAC TTCGCCCCAC CCGTCGGCGG
   351  CAGGGAAGCC CGCyTGAGCC ACGACGAAAT CCTCATCCTC TTCCACGAAA
   401  CCGGACACGG GCTGCACCAC CTGCTTACCC AAGTGGACGA ACTGGGCGTA
   451  TCCGGCATCA ACGGCGTAkA ATGGGACGCG GTCGAACTGC CCAGCCAGTT
   501  TATGGAAAAT TTCGTTTGGG AATACAATGT CTTGGCACAA mTGTCAGCCC
   551  ACGAAGAAAC CGGcgTTCCC yTGCCGAAAG AACTCTTsGA CAAAwTGCTC
   601  GCCGCCAAAA ACTTCCAAsG CGGCATGTTC yTsGTCCGGC AAwTGGAGTT
   651  CGCCCTCTTT GATATGATGA TTTACAGCGA AGACGACGAA GGCCGTCTGA
   701  AAAACTGGCA ACAGGTTTTA GACAGCGTGC GCAAAAAAGT CGCCGTCATC
   751  CAGCCGCCCG AATACAACCG CTTCGCCTTG AGCTTCGGCC ACATCTTCGC
   801  AGGCGGCTAT TCCGCAGCTn ATTACAGCTA CGCGTGGGCG GAAGTATTGA
   851  GCGCGGACGC ATACGCCGCC TTTGAAGAAA GCGACGATGT CGCCGCCACA
   901  GGCAAACGCT TTTGGCAGGA AATCCTCGCC GTCGGGGnAT CGCGCAGCGG
   951  nGCAGAATCC TTCAAAGCCT TCCGCGGCCG CGAACCGAGC ATAGACGCAC
  1001  TCTTGCGCCA CAGCGGTTTC GACAACGCGG TCTGA
```

This corresponds to the amino acid sequence <SEQ ID 3086; ORF 128>:

```
m128.pep  (partial)
     1  MTDNALLHLG EEPRFDQIKT EDIKPALQTA IAEAREQIAA IKAQTHTGWA
    51  NTVEPLTGIT ERVGRIWGVV SHLNCVADTP ELRAVYNELM PEITVFFTEI
   101  GQDIELYNRF KTIKNSPEFD TLSPAQKTKL NH
//
     1  YASEKLREAK YAFSETXVKK YFPVGXVLNG LFAQXKKLYG IGFTEKTVPV
    51  WHKDVRYXEL QQNGEXIGGV YMDLYAREGK RGGAWMNDYK GRRRFSDGTL
   101  QLPTAYLVCN FAPPVGGREA RLSHDEILIL FHETGHGLHH LLTQVDELGV
   151  SGINGVXWDA VELPSQFMEN FVWEYNVLAQ XSAHEETGVP LPKELXDKXL
   201  AAKNFQXGMF XVRQXEFALF DMMIYSEDDE GRLKNWQQVL DSVRKKVAVI
   251  QPPEYNRFAL SFGHIFAGGY SAAXYSYAWA EVLSADAYAA FEESDDVAAT
   301  GKRFWQEILA VGXSRSGAES FKAFRGREPS IDALLRHSGF DNAV*
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3087>:

```
g128.seq
   1  atgattgaca acgCActgct ccacttgggc gaagaaccCC GTTTTaatca
  51  aatccaaacc gaagACAtca AACCCGCCGT CCAAACCGCC ATCGCCGAAG
 101  CGCGCGGACA AATCGCCGCC GTCAAAGCGC AAACGCACAC CGGCTGGGCG
 151  AACACCGTCG AGCGTCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
 201  GGGCGTCGTG TCCCATCTCA ACTCCGTCGT CGACACGCCC GAACTGCGCG
 251  CCGTCTATAA CGAACTGATG CCTGAAATCA CCGTCTTCTT CACCGAAATC
 301  GGACAAGACA TCGAACTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
 351  CGAATTTGCA ACGCTTTCCC CCGCACAAAA AACCAAGCTC GATCACGACC
 401  TGCGCGATTT CGTATTGAGC GGCGCGGAAC TGCCGCCCGA ACGGCAGGCA
 451  GAACTGGCAA AACTGCAAAC CGAAGGCGCG CAACTTTCCG CCAAATTCTC
 501  CCAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTTAC TTTGACGATG
 551  CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCC
 601  GCCGCGCAAA GCGAAGGCAA AACAGGTTAC AAAATCGGCT TGCAGATTCC
 651  GCACTACCTT GCCGTTATCC AATACGCCGG CAACCGCGAA CTGCGCGAAC
 701  AAATCTACCG CGCCTACGTT ACCCGTGCCA GCGAACTTTC AAACGACGGC
 751  AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCATTGAA
 801  AACCGccaaa cTGCTCGGCT TTAAAAATTA CGCCGAATTG TCGCTGGCAA
 851  CCAAAATGGC GGACACGCCC GAACAGGTTT TAAACTTCCT GCACGACCTC
 901  GCCCGCCGCG CCAAACCCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
 951  CTTCGCCCGC GAACACCTCG GTCTCGCCGA CCCGCAGCCG TGGGACTTGA
1001  GCTACGCCGG CGAAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
1051  GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTTCTGGCAG GCCTGTTCGC
1101  CCAAATCAAA AAACTCTACG GCATCGGATT CGCCGAAAAA ACCGTTCCCG
1151  TCTGGCACAA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGGCAAAACC
1201  ATCGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
1251  CGCGTGGATG AACGACtaca AAGGCCGCCG CCGCTTTGCC GACGgcacGC
1301  TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCGCCCC GCCCGTCGGC
1351  GGCAAAGAAG CGCGTTTAAG CCACGACGAA ATCCTCACCC TCTTCCACGA
1401  AacCGGCCAC GGACTGCACC ACCTGCTTAC CCAAGTGGAC GAACTGGGCG
1451  TGTCCGGCAT CAAcggcgtA GAATGGGACG CGGTCGAACT GCCCAGCCAG
1501  TTTATGGAAA ACTTCGTTTG GGAATACAAT GTATTGGCAC AAATGTCCGC
1551  CCACGAAGAA AccgGCGAGC CCCTGCCGAA AGAACTCTTC GACAAAATGC
1601  TcgcCGCCAA AAACTTCCAG CGCGGTATGT TCCTCGTCCG GCAAATGGAG
1651  TTCGCCCTCT TCGATATGAT GATTTACAGT GAAAGCGACG AATGCCGTCT
1701  GAAAAACTGG CAGCAGGTTT TAGACAGCGT GCGCAAAGAA GTcGCCGTCA
1751  TCCAACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCacatctTC
1801  GCcggcGGCT ATTCCGCAGG CTATTACAGC TACGCATGGG CCGAAGTCCt
1851  cAGCACCGAT GCCTACGCCG CCTTTGAAGA AAGcGACGac gtcGCCGCCA
1901  CAGGCAAACG CTTCTGGCAA GAAAtccttg ccgtcggcgg ctCCCGCAGC
1951  gcgGCGGAAT CCTTCAAAGC CTTCCGCGGA CGCGAACCGA GCATAGACGC
2001  ACTGCTGCGC CAaagcggtT TCGACAACGC gGCttgA
```

This corresponds to the amino acid sequence <SEQ ID 3088; ORF 128.ng>:

```
g128.pep
   1  MIDNALLHLG EEPRFNQIQT EDIKPAVQTA IAEARGQIAA VKAQTHTGWA
  51  NTVERLTGIT ERVGRIWGVV SHLNSVVDTP ELRAVYNELM PEITVFFTEI
 101  GQDIELYNRF KTIKNSPEFA TLSPAQKTKL DHDLRDFVLS GAELPPERQA
 151  ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
 201  AAQSEGKTGY KIGLQIPHYL AVIQYAGNRE LREQIYRAYV TRASELSNDG
 251  KFDNTANIDR TLENALKTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
 301  ARRAKPYAEK DLAEVKAFAR EHLGLADPQP WDLSYAGEKL REAKYAFSET
 351  EVKKYFPVGK VLAGLFAQIK KLYGIGFAEK TVPVWHKDVR YFELQQNGKT
 401  IGGVYMDLYA REGKRGGAWM NDYKGRRRFA DGTLQLPTAY LVCNFAPPVG
 451  GKEARLSHDE ILTLFHETGH GLHHLLTQVD ELGVSGINGV EWDAVELPSQ
 501  FMENFVWEYN VLAQMSAHEE TGEPLPKELF DKMLAAKNFQ RGMFLVRQME
 551  FALFDMMIYS ESDECRLKNW QQVLDSVRKE VAVIQPPEYN RFANSFGHIF
 601  AGGYSAGYYS YAWAEVLSTD AYAAFEESDD VAATGKRFWQ EILAVGGSRS
```

```
651  AAESFKAFRG REPSIDALLR QSGFDNAA*
```

ORF 128 shows 91.7% identity over a 475 aa overlap with a predicted ORF (ORF 128.ng) from *N. gonorrhoeae:*

m128/g128

```
                      10        20        30        40        50        60
g128.pep      MIDNALLHLGEEPRFNQIQTEDIKPAVQTAIAEARGQIAAVKAQTHTGWANTVERLTGIT
              | ||||||||||||:||:||||||:|||||||| ||||:|||||||||||| |||||
m128          MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIKAQTHTGWANTVEPLTGIT
                      10        20        30        40        50        60

                      70        80        90       100       110       120
g128.pep      ERVGRIWGVVSHLNSVVDTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFA
              ||||||||||||||  |:||||||||||||||||||||||||||||||||||||||||||
m128          ERVGRIWGVVSHLNCVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
                      70        80        90       100       110       120

                     130       140       150       160       170       180
g128.pep      TLSPAQKTKLDHDLRDFVLSGAELPPERQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY
              ||||||||||:|
m128          TLSPAQKTKLNH
                     130
       //
                                                 340       350       360
g128.pep                                  YAGEKLREAKYAFSETEVKKYFPVGKVLAG
                                          ||:||||||||||||| |||||||| || |
m128                                      YASEKLREAKYAFSETXVKKYFPVGXVLNG
                                                  10        20        30

                     370       380       390       400       410       420
g128.pep      LFAQIKKLYGIGFAEKTVPVWHKDVRYFELQQNGKTIGGVYMDLYAREGKRGGAWMNDYK
              |||| ||||||||:|||||||||||| ||||||::||||||||||||||||||||||||
m128          LFAQXKKLYGIGFTEKTVPVWHKDVRYXELQQNGEXIGGVYMDLYAREGKRGGAWMNDYK
                      40        50        60        70        80        90

                     430       440       450       460       470       480
g128.pep      GRRRFADGTLQLPTAYLVCNFAPPVGGKEARLSHDEILTLFHETGHGLHHLLTQVDELGV
              |||||:|||||||||||||||||||||:||||||||||| ||||||||||||||||||||
m128          GRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILILFHETGHGLHHLLTQVDELGV
                     100       110       120       130       140       150

                     490       500       510       520       530       540
g128.pep      SGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGEPLPKELFDKMLAAKNFQRGMF
              |||||| ||||||||||||||||||||||| ||||||| |||||| || ||||||| |||
m128          SGINGVXWDAVELPSQFMENFVWEYNVLAQXSAHEETGVPLPKELXDKXLAAKNFQXGMF
                     160       170       180       190       200       210

                     550       560       570       580       590       600
g128.pep      LVRQMEFALFDMMIYSESDECRLKNWQQVLDSVRKEVAVIQPPEYNRFANSFGHIFAGGY
               ||| ||||||||||||:|| |||||||||||||:|||||||||||| |||||||||
m128          XVRQXEFALFDMMIYSEDDEGRLKNWQQVLDSVRKKVAVIQPPEYNRFALSFGHIFAGGY
                     220       230       240       250       260       270

                     610       620       630       640       650       660
g128.pep      SAGYYSYAWAEVLSTDAYAAFEESDDVAATGKRFWQEILAVGGSRSAAESFKAFRGREPS
              ||: ||||||||||:||||||||||||||||||||||||||| |||:|||||||||||||
m128          SAAXYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGXSRSGAESFKAFRGREPS
                     280       290       300       310       320       330
```

```
                670       679
g128.pep     IDALLRQSGFDNAAX
             ||||||:||||||:
m128         IDALLRHSGFDNAVX
                    340
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3089>:

a128.seq

1 ATGACTGACA ACGCACTGCT CCATTTGGGC GAAGAACCCC GTTTTGATCA
51 AATCAAAACC GAAGACATCA AACCCGCCCT GCAAACCGCC ATTGCCGAAG
101 CGCGCGAACA AATCGCCGCC ATCAAAGCCC AAACGCACAC CGGCTGGGCA
151 AACACTGTCG AACCCCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
201 GGGCGTGGTG TCGCACCTCA ACTCCGTCAC CGACACGCCC GAACTGCGCG
251 CCGCCTACAA TGAATTAATG CCCGAAATTA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAACTCCCC
351 CGAGTTCGAC ACCCTCTCCC ACGCGCAAAA AACCAAACTC AACCACGATC
401 TGCGCGATTT CGTCCTCAGC GGCGCGGAAC TGCCGCCCGA ACAGCAGGCA
451 GAATTGGCAA AACTGCAAAC CGAAGGCGCG CAACTTTCCG CCAAATTCTC
501 CCAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTTAC TTTGACGATG
551 CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCT
601 GCCGCGCAAA GCGAAGGCAA AACAGGCTAC AAAATCGGTT TGCAGATTCC
651 GCACTACCTC GCCGTCATCC AATACGCCGA [illegible]CGCAAA CTGCGCGAAC
701 AAATCTACCG CGCCTACGTT ACCCGCGCCA GCGAGCTTTC AGACGACGGC
751 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCCCTGCA
801 AACCGCCAAA CTGCTCGGCT TCAAAAACTA CGCCGAATTG TCGCTGGCAA
851 CCAAAATGGC GGACACCCCC GAACAAGTTT TAAACTTCCT GCACGACCTC
901 GCCCGCCGCG CCAAACCCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
951 CTTCGCCCGC GAAAGCCTCG GCCTCGCCGA TTTGCAACCG TGGGACTTGG
1001 GCTACGCCGG CGAAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTATTAAACG GACTGTTCGC
1101 CCAAATCAAA AAACTCTACG GCATCGGATT TACCGAAAAA ACCGTCCCCG
1151 TCTGGCACAA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGGCGAAACC
1201 ATAGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
1251 CGCGTGGATG AACGACTACA AAGGCCGCCG CCGTTTTTCA GACGGCACGC
1301 TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCACCCC GCCCGTCGGC
1351 GGCAAAGAAG CCCGCTTGAG CCATGACGAA ATCCTCACCC TCTTCCACGA
1401 AACCGGACAC GGCCTGCACC ACCTGCTTAC CCAAGTCGAC GAACTGGGCG
1451 TATCCGGCAT CAACGGCGTA GAATGGGACG CAGTCGAACT GCCCAGTCAG
1501 TTTATGGAAA ATTTCGTTTG GGAATACAAT GTCTTGGCGC AAATGTCCGC
1551 CCACGAAGAA ACCGGCGTTC CCCTGCCGAA AGAACTCTTC GACAAAATGC
1601 TCGCCGCCAA AAACTTCCAA CGCGGAATGT TCCTCGTCCG CCAAATGGAG
1651 TTCGCCCTCT TTGATATGAT GATTTACAGC GAAGACGACG AAGGCCGTCT
1701 GAAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAGAA GTCGCCGTCG
1751 TCCGACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCACATCTTC
1801 GCAGGCGGCT ATTCCGCAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
1851 GAGCGCGGAC GCATACGCCG CCTTTGAAGA AAGCGACGAT GTCGCCGCCA
1901 CAGGCAAACG CTTTTGGCAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC
1951 GCGGCAGAAT CCTTCAAAGC CTTCCGCGGA CGCGAACCGA GCATAGACGC
2001 ACTCTTGCGC CACAGCGGCT TCGACAACGC GGCTTGA

This corresponds to the amino acid sequence <SEQ ID 3090; ORF 128.a>:

a128.pep

1 MTDNALLHLG EEPRFDQIKT EDIKPALQTA IAEAREQIAA IKAQTHTGWA
51 NTVEPLTGIT ERVGRIWGVV SHLNSVTDTP ELRAAYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFD TLSHAQKTKL NHDLRDFVLS GAELPPEQQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSEGKTGY KIGLQIPHYL AVIQYADNRK LREQIYRAYV TRASELSDDG
251 KFDNTANIDR TLENALQTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
301 ARRAKPYAEK DLAEVKAFAR ESLGLADLQP WDLGYAGEKL REAKYAFSET
351 EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVWHKDVR YFELQQNGET
401 IGGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTLQLPTAY LVCNFTPPVG
451 GKEARLSHDE ILTLFHETGH GLHHLLTQVD ELGVSGINGV EWDAVELPSQ
501 FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME

```
551  FALFDMMIYS EDDEGRLKNW QQVLDSVRKE VAVVRPPEYN RFANSFGHIF
601  AGGYSAGYYS YAWA_.LSAD AYAAFEESDD VAATGKRFWQ EILAVGGSRS
651  AAESFKAFRG REPSIDALLR HSGFDNAA*
```

m128/a128 ORFs 128 and 128.a showed a 66.0% identity in 677 aa overlap

```
                      10        20        30        40        50        60
m128.pep    M:DNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIKAQTHTGWANTVEPLTGIT
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a128        MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIKAQTHTGWANTVEPLTGIT
                      10        20        30        40        50        60

                      70        80        90       100       110       120
m128.pep    ERVGRIWGVVSHLNCVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
            |||||||||||||| |:|||||||:|||||||||||||||||||||||||||||||||||
a128        ERVGRIWGVVSHLNSVTDTPELRAAYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
                      70        80        90       100       110       120

                     130
m128.pep    TLSPAQKTKLNH------------------------------------------------
            ||| ||||||||
a128        TLSHAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY
                     130       140       150       160       170       180

m128.pep    ------------------------------------------------------------

a128        FDDAAPLAGIPEDALAMFAAAAQSEGKTGYKIGLQIPHYLAVIQYADNRKLREQIYRAYV
                     190       200       210       220       230       240

m128.pep    ------------------------------------------------------------

a128        TRASELSDDGKFDNTANIDRTLENALQTAKLLGFKNYAELSLATKMADTPEQVLNFLHDL
                     250       260       270       280       290       300

                                                      140       150
m128.pep    ----------------------------------YASEKLREAKYAFSETXVKKYFPVGX
                                              ||:||||||||||||| ||||||||
a128        ARRAKPYAEKDLAEVKAFARESLGLADLQPWDLGYAGEKLREAKYAFSETEVKKYFPVGK
                     310       320       330       340       350       360

           160       170       180       190       200       210
m128.pep    VLNGLFAQXKKLYGIGFTEKTVPVWHKDVRYXELQQNGEXIGGVYMDLYAREGKRGGAWM
            |||||||| |||||||||||||||||||||| ||||||| :||||||||||||||||||||
a128        VLNGLFAQIKKLYGIGFTEKTVPVWHKDVRYFELQQNGETIGGVYMDLYAREGKRGGAWM
                     370       380       390       400       410       420

           220       230       240       250       260       270
m128.pep    NDYKGRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILILFHETGHGLHHLLTQVD
            |||||||||||||||||||||||||:|||||:|||||||||| |||||||||||||||||
a128        NDYKGRRRFSDGTLQLPTAYLVCNFTPPVGGKEARLSHDEILTLFHETGHGLHHLLTQVD
                     430       440       450       460       470       480

           280       290       300       310       320       330
m128.pep    ELGVSGINGVXWDAVELPSQFMENFVWEYNVLAQXSAHEETGVPLPKELXDKXLAAKNFQ
            |||||||||| ||||||||||||||||||||||| |||||||||||||| || |||||||
a128        ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ
                     490       500       510       520       530       540

           340       350       360       370       380       390
m128.pep    XGMFXVRQXEFALFDMMIYSEDDEGRLKNWQQVLDSVRKKVAVIQPPEYNRFALSFGHIF
             ||| ||| ||||||||||||||||||||||||||||||:|||::|||||||| ||||||
a128        RGMFLVRQMEFALFDMMIYSEDDEGRLKNWQQVLDSVRKEVAVVRPPEYNRFANSFGHIF
                     550       560       570       580       590       600
```

```
                400       410       420       430       440       450
m128.pep    AGGYSAAXYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGXSRSGAESFKAFRG
            ||||||: ||||||||||||||||||||||||||||||||||||||| |||:||||||||
a128        AGGYSAGYYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGGSRSAAESFKAFRG
                     610       620       630       640       650       660

                460       470
m128.pep    REPSIDALLRHSGFDNAVX
            |||||||||||||||||:
a128        REPSIDALLRHSGFDNAAX
                     670
```

Further work revealed the DNA sequence identified in *N. meningitidis* <SEQ ID 3091>:

```
m128-1.seq
     1  ATGACTGACA ACGCACTGCT CCATTTGGGC GAAGAACCCC GTTTTGATCA
    51  AATCAAAACC GAAGACATCA AACCCGCCC. GCAAACCGCC ATCGCCGAAG
   101  CGCGCGAACA AATCGCCGCC ATCAAAGCCC AAACGCACAC CGGCTGGGCA
   151  AACACTGTCG AACCCCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
   201  GGGCGTGGTG TCGCACCTCA ACTCCGTCGC CGACACGCCC GAACTGCGCG
   251  CCGTCTATAA CGAACTGATG CCCGAAATCA CCGTCTTCTT CACCGAAATC
   301  GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
   351  CGAATTCGAC ACCCTCTCCC CCGCACAAAA AACCAAACTC AACCACGATC
   401  TGCGCGATTT CGTCCTCAGC GGCGCGGAAC TGCCGCCCGA ACAGCAGGCA
   451  GAACTGGCAA AACTGCAAAC CGAAGGCGCG CAACTTTCCG CCAAATTCTC
   501  CCAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTTAC TTTGACGATG
   551  CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCC
   601  GCCGCGCAAA GCGAAAGCAA AACAGGCTAC AAAATCGGCT TGCAGATTCC
   651  ACACTACCTC GCCGTCATCC AATACGCCGA CAACCGCGAA CTGCGCGAAC
   701  AAATCTACCG CGCCTACGTT ACCCGCGCCA GCGAACTTTC AGACGACGGC
   751  AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGCAA ACGCCCTGCA
   801  AACCGCCAAA CTGCTCGGCT TCAAAAACTA CGCCGAATTG TCGCTGGCAA
   851  CCAAAATGGC GGACACGCCC GAACAAGTTT TAAACTTCCT GCACGACCTC
   901  GCCCGCCGCG CCAAACCCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
   951  CTTCGCCCGC GAAAGCCTGA ACCTCGCCGA TTTGCAACCG TGGGACTTGG
  1001  GCTACGCCAG CGAAAAACTG CGCGAAGCCA AATACGCGTT CAGCGAAACC
  1051  GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTATTAAACG GACTGTTCGC
  1101  CCAAATCAAA AAACTCTACG GCATCGGATT TACCGAAAAA ACCGTCCCCG
  1151  TCTGGCACAA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGGCGAAACC
  1201  ATAGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
  1251  CGCGTGGATG AACGACTACA AAGGCCGCCG CCGTTTTTCA GACGGCACGC
  1301  TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCGCCCC ACCCGTCGGC
  1351  GGCAGGGAAG CCCGCCTGAG CCACGACGAA ATCCTCATCC TCTTCCACGA
  1401  AACCGGACAC GGGCTGCACC ACCTGCTTAC CCAAGTGGAC GAACTGGGCG
  1451  TATCCGGCAT CAACGGCGTA GAATGGGACG CGGTCGAACT GCCCAGCCAG
  1501  TTTATGGAAA ATTTCGTTTG GGAATACAAT GTCTTGGCAC AAATGTCAGC
  1551  CCACGAAGAA ACCGGCGTTC CCCTGCCGAA AGAACTCTTC GACAAAATGC
  1601  TCGCCGCCAA AAACTTCCAA CGCGGCATGT TCCTCGTCCG GCAAATGGAG
  1651  TTCGCCCTCT TTGATATGAT GATTTACAGC GAAGACGACG AAGGCCGTCT
  1701  GAAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAAAA GTCGCCGTCA
  1751  TCCAGCCGCC CGAATACAAC CGCTTCGCCT TGAGCTTCGG CCACATCTTC
  1801  GCAGGCGGCT ATTCCGCAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
  1851  GAGCGCGGAC GCATACGCCG CCTTTGAAGA AAGCGACGAT GTCGCCGCCA
  1901  CAGGCAAACG CTTTTGGCAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC
  1951  GCGGCAGAAT CCTTCAAAGC CTTCCGCGGC CGCGAACCGA GCATAGACGC
  2001  ACTCTTGCGC CACAGCGGTT TCGACAACGC GGTCTGA
```

This corresponds to the amino acid sequence <SEQ ID 3092; ORF 128-1>:

```
m128-1.pep.
     1  MTDNALLHLG EEPRFDQIKT EDIKPALQTA IAEAREQIAA IKAQTHTGWA
    51  NTVEPLTGIT ERVGRIWGVV SHLNSVADTP ELRAVYNELM PEITVFFTEI
```

```
101  GQDIELYNRF KTIKNSPEFD TLSPAQKTKL NHDLRDFVLS GAELPPEQQA
151  ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201  AAQSESKTGY KIGLQIPHYL AVIQYADNRE LREQIYRAYV TRASELSDDG
251  KFDNTANIDR TLANALQTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
301  ARRAKPYAEK DLAEVKAFAR ESLNLADLQP WDLGYASEKL REAKYAFSET
351  EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVWHKDVR YFELQQNGET
401  IGGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTLQLPTAY LVCNFAPPVG
451  GREARLSHDE ILILFHETGH GLHHLLTQVD ELGVSGINGV EWDAVELPSQ
501  FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME
551  FALFDMMIYS EDDEGRLKNW QQVLDSVRKK VAVIQPPEYN RFALSFGHIF
601  AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRFWQ EILAVGGSRS
651  AAESFKAFRG REPSIDALLR HSGFDNAV*
```

The following DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3093>:

```
g128-1.seq  (partial)
   1  ATGATTGACA ACGCACTGCT CCACTTGGGC GAAGAACCCC GTTTTAATCA
  51  AATCAAAACC GAAGACATCA AACCCGCCGT CCAAACCGCC ATCGCCGAAG
 101  CGCGCGGACA AATCGCCGCC GTCAAAGCGC AAACGCACAC CGGCTGGGCG
 151  AACACCGTCG AGCGTCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
 201  GGGCGTCGTG TCCCATCTCA ACTCCGTCGT CGACACGCCC GAACTGCGCG
 251  CCGTCTATAA CGAACTGATG CCTGAAATCA CCGTCTTCTT CACCGAAATC
 301  GGACAAGACA TCGAACTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
 351  CGAATTTGCA ACGCTTTCCC CCGCACAAAA AACCAAGCTC GATCACGACC
 401  TGCGCGATTT CGTATTGAGC GGCGCGGAAC TGCCGCCCGA ACGGCAGGCA
 451  GAACTGGCAA AACTGCAAAC CGAAGGCGCG CAACTTTCCG CCAAATTCTC
 501  CCAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTTAC TTTGACGATG
 551  CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCC
 601  GCCGCGCAAA GCGAAGGCAA AACAGGTTAC AAAATCGGCT TGCAGATTCC
 651  GCACTACCTT GCCGTTATCC AATACGCCGG CAACCGCGAA CTGCGCGAAC
 701  AAATCTACCG CGCCTACGTT ACCCGTGCCA GCGAACTTTC AAACGACGGC
 751  AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCATTGAA
 801  AACCGCCAAA CTGCTCGGCT TTAAAAATTA CGCCGAATTG TCGCTGGCAA
 851  CCAAAATGGC GGACACGCCC GAACAGGTTT TAAACTTCCT GCACGACCTC
 901  GCCCGCCGCG CCAAACCCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
 951  CTTCGCCCGC GAACACCTCG GTCTCGCCGA CCCGCAGCCG TGGGACTTGA
1001  GCTACGCCGG CGAAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
1051  GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTTCTGGCAG GCCTGTTCGC
1101  CCAAATCAAA AAACTCTACG GCATCGGATT CGCCGAAAAA ACCGTTCCCG
1151  TCTGGCACAA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGGCAAAACC
1201  ATCGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
1251  CGCGTGGATG AACGACTACA AAGGCCGCCG CCGCTTTGCC GACGGCACGC
1301  TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCGCCCC GCCCGTCGGC
1351  GGCAAAGAAG CGCGTTTAAG CCACGACGAA ATCCTCACCC TCTTCCACGA
1401  AACCGGCCAC GGACTGCACC ACCTGCTTAC CCAAGTGGAC GAACTGGGCG
1451  TGTCCGGCAT CAACGGCGTA AAA
```

This corresponds to the amino acid sequence <SEQ ID 3094; ORF 128-1.ng>:

```
g128-1.pep  (partial)
   1  MIDNALLHLG EEPRFNQIKT EDIKPAVQTA IAEARGQIAA VKAQTHTGWA
  51  NTVERLTGIT ERVGRIWGVV SHLNSVVDTP ELRAVYNELM PEITVFFTEI
 101  GQDIELYNRF KTIKNSPEFA TLSPAQKTKL DHDLRDFVLS GAELPPERQA
 151  ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
 201  AAQSEGKTGY KIGLQIPHYL AVIQYAGNRE LREQIYRAYV TRASELSNDG
 251  KFDNTANIDR TLENALKTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
 301  ARRAKPYAEK DLAEVKAFAR EHLGLADPQP WDLSYAGEKL REAKYAFSET
 351  EVKKYFPVGK VLAGLFAQIK KLYGIGFAEK TVPVWHKDVR YFELQQNGKT
 401  IGGVYMDLYA REGKRGGAWM NDYKGRRRFA DGTLQLPTAY LVCNFAPPVG
 451  GKEARLSHDE ILTLFHETGH GLHHLLTQVD ELGVSGINGV K
```

m128-1/g128-1 ORFs 128-1 and 128-1.ng showed a 94.5% identity in 491 aa overlap

```
            10        20        30        40        50        60
```

```
g128-1.pep        MIDNALLHLGEEPRFNQIKTEDIKPAVQTAIAEARGQIAAVKAQTHTGWANTVERLTGIT
                  | |||||||||||||:||||||||||:||||||||  ||||:||||||||||||| |||||
m128-1            MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIKAQTHTGWANTVEPLTGIT
                          10        20        30        40        50        60

                          70        80        90       100       110       120
g128-1.pep        ERVGRIWGVVSHLNSVVDTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFA
                  ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
m128-1            ERVGRIWGVVSHLNSVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
                          70        80        90       100       110       120

                         130       140       150       160       170       180
g128-1.pep        TLSPAQKTKLDHDLRDFVLSGAELPPERQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY
                  ||||||||||:||||||||||||||||:||||||||||||||||||||||||||||||||
m128-1            TLSPAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY
                         130       140       150       160       170       180

                         190       200       210       220       230       240
g128-1.pep        FDDAAPLAGIPEDALAMFAAAAQSEGKTGYKIGLQIPHYLAVIQYAGNRELREQIYRAYV
                  |||||||||||||||||||||||||:|||||||||||||||||||| |||||||||||||
m128-1            FDDAAPLAGIPEDALAMFAAAAQSESKTGYKIGLQIPHYLAVIQYADNRELREQIYRAYV
                         190       200       210       220       230       240

                         250       260       270       280       290       300
g128-1.pep        TRASELSNDGKFDNTANIDRTLENALKTAKLLGFKNYAELSLATKMADTPEQVLNFLHDL
                  |||||||:||||||||||||||| ||:|||||||||||||||||||||||||||||||||
m128-1            TRASELSDDGKFDNTANIDRTLANALQTAKLLGFKNYAELSLATKMADTPEQVLNFLHDL
                         250       260       270       280       290       300

                         310       320       330       340       350       360
g128-1.pep        ARRAKPYAEKDLAEVKAFAREHLGLADPQPWDLSYAGEKLREAKYAFSETEVKKYFPVGK
                  ||||||||||||||||||||| |:||| |||||:||:|||||||||||||||||||||||
m128-1            ARRAKPYAEKDLAEVKAFARESLNLADLQPWDLGYASEKLREAKYAFSETEVKKYFPVGK
                         310       320       330       340       350       360

                         370       380       390       400       410       420
g128-1.pep        VLAGLFAQIKKLYGIGFAEKTVPVWHKDVRYFELQQNGKTIGGVYMDLYAREGKRGGAWM
                  || ||||||||||||||:|||||||||||||||||||||:||||||||||||||||||||
m128-1            VLNGLFAQIKKLYGIGFTEKTVPVWHKDVRYFELQQNGETIGGVYMDLYAREGKRGGAWM
                         370       380       390       400       410       420

                         430       440       450       460       470       480
g128-1.pep        NDYKGRRRFADGTLQLPTAYLVCNFAPPVGGKEARLSHDEILTLFHETGHGLHHLLTQVD
                  |||||||||:|||||||||||||||||||||:|||||||||| |||||||||||||||||
m128-1            NDYKGRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILILFHETGHGLHHLLTQVD
                         430       440       450       460       470       480

                         490
g128-1.pep        ELGVSGINGVK
                  ||||||||||:
m128-1            ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ
                         490       500       510       520       530       540
```

The following DNA sequence was identified in *N. meningitidis* <SEQ ID 3095>:

```
a128-1.seq
      1  ATGACTGACA ACGCACTGCT CCATTTGGGC GAAGAACCCC GTTTTGATCA
     51  AATCAAAACC GAAGACATCA AACCCGCCCT GCAAACCGCC ATTGCCGAAG
    101  CGCGCGAACA AATCGCCGCC ATCAAAGCCC AAACGCACAC CGGCTGGGCA
    151  AACACTGTCG AACCCCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
    201  GGGCGTGGTG TCGCACCTCA ACTCCGTCAC CGACACGCCC GAACTGCGCG
    251  CCGCCTACAA TGAATTAATG CCCGAAATTA CCGTCTTCTT CACCGAAATC
    301  GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAACTCCCC
    351  CGAGTTCGAC ACCCTCTCCC ACGCGCAAAA AACCAAACTC AACCACGATC
    401  TGCGCGATTT CGTCCTCAGC GGCGCGGAAC TGCCGCCCGA ACAGCAGGCA
    451  GAATTGGCAA AACTGCAAAC CGAAGGCGCG CAACTTTCCG CCAAATTCTC
```

```
 501  CCAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTTAC TTTGACGATG
 551  CCGCACCGCT TGCC..CATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCT
 601  GCCGCGCAAA GCGAAGGCAA AACAGGCTAC AAAATCGGTT TGCAGATTCC
 651  GCACTACCTC GCCGTCATCC AATACGCCGA CAACCGCAAA CTGCGCGAAC
 701  AAATCTACCG CGCCTACGTT ACCCGCGCCA GCGAGCTTTC AGACGACGGC
 751  AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCCCTGCA
 801  AACCGCCAAA CTGCTCGGCT TCAAAAACTA CGCCGAATTG TCGCTGGCAA
 851  CCAAAATGGC GGACACCCCC GAACAAGTTT TAAACTTCCT GCACGACCTC
 901  GCCCGCCGCG CCAAACCCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
 951  CTTCGCCCGC GAAAGCCTCG GCCTCGCCGA TTTGCAACCG TGGGACTTGG
1001  GCTACGCCGG CGAAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
1051  GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTATTAAACG GACTGTTCGC
1101  CCAAATCAAA AAACTCTACG GCATCGGATT TACCGAAAAA ACCGTCCCCG
1151  TCTGGCACAA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGGCGAAACC
1201  ATAGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
1251  CGCGTGGATG AACGACTACA AAGGCCGCCG CCGTTTTTCA GACGGCACGC
1301  TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCACCCC GCCCGTCGGC
1351  GGCAAAGAAG CCCGCTTGAG CCATGACGAA ATCCTCACCC TCTTCCACGA
1401  AACCGGACAC GGCCTGCACC ACCTGCTTAC CCAAGTCGAC GAACTGGGCG
1451  TATCCGGCAT CAACGGCGTA GAATGGGACG CAGTCGAACT GCCCAGTCAG
1501  TTTATGGAAA ATTTCGTTTG GGAATACAAT GTCTTGGCGC AAATGTCCGC
1551  CCACGAAGAA ACCGGCGTTC CCCTGCCGAA AGAACTCTTC GACAAAATGC
1601  TCGCCGCCAA AAACTTCCAA CGCGGAATGT TCCTCGTCCG CCAAATGGAG
1651  TTCGCCCTCT TTGATATGAT GATTTACAGC GAAGACGACG AAGGCCGTCT
1701  GAAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAGAA GTCGCCGTCG
1751  TCCGACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCACATCTTC
1801  GCAGGCGGCT ATTCCGCAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
1851  GAGCGCGGAC GCATACGCCG CCTTTGAAGA AAGCGACGAT GTCGCCGCCA
1901  CAGGCAAACG CTTTTGGCAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC
1951  GCGGCAGAAT CCTTCAAAGC CTTCCGCGGA CGCGAACCGA GCATAGACGC
2001  ACTCTTGCGC CACAGCGGCT TCGACAACGC GGCTTGA
```

This corresponds to the amino acid sequence <SEQ ID 3096; ORF 128-1.a>:

```
a128-1.pep
  1  MTDNALLHLG EEPRFDQIKT EDIKPALQTA IAEAREQIAA IKAQTHTGWA
 51  NTVEPLTGIT ERVGRIWGVV SHLNSVTDTP ELRAAYNELM PEITVFFTEI
101  GQDIELYNRF KTIKNSPEFD TLSHAQKTKL NHDLRDFVLS GAELPPEQQA
151  ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201  AAQSEGKTGY KIGLQIPHYL AVIQYADNRK LREQIYRAYV TRASELSDDG
251  KFDNTANIDR TLENALQTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
301  ARRAKPYAEK DLAEVKAFAR ESLGLADLQP WDLGYAGEKL REAKYAFSET
351  EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVWHKDVR YFELQQNGET
401  IGGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTLQLPTAY LVCNFTPPVG
451  GKEARLSHDE ILTLFHETGH GLHHLLTQVD ELGVSGINGV EWDAVELPSQ
501  FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME
551  FALFDMMIYS EDDEGRLKNW QQVLDSVRKE VAVVRPPEYN RFANSFGHIF
601  AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRFWQ EILAVGGSRS
651  AAESFKAFRG REPSIDALLR HSGFDNAA*
```

m128-1/a128-1 ORFs 128-1 and 128-1.a showed a 97.8% identity in 677 aa overlap

```
                    10        20        30        40        50        60
a128-1.pep  MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIKAQTHTGWANTVEPLTGIT
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m128-1      MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIKAQTHTGWANTVEPLTGIT
                    10        20        30        40        50        60

                    70        80        90       100       110       120
a128-1.pep  ERVGRIWGVVSHLNSVTDTPELRAAYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
            ||||||||||||||||:||||||:||||||||||||||||||||||||||||||||||||
m128-1      ERVGRIWGVVSHLNSVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
                    70        80        90       100       110       120

                   130       140       150       160       170       180
```

```
a128-1.pep    TLSHAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m128-1        TLSPAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY
                     130       140       150       160       170       180

                     190       200       210       220       230       240
a128-1.pep    FDDAAPLAGIPEDALAMFAAAAQSEGKTGYKIGLQIPHYLAVIQYADNRKLREQIYRAYV
              |||||||||||||||||||||||||:|||||||||||||||||||||||:||||||||||
m128-1        FDDAAPLAGIPEDALAMFAAAAQSESKTGYKIGLQIPHYLAVIQYADNRELREQIYRAYV
                     190       200       210       220       230       240

                     250       260       270       280       290       300
a128-1.pep    TRASELSDDGKFDNTANIDRTLENALQTAKLLGFKNYAELSLATKMADTPEQVLNFLHDL
              |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
m128-1        TRASELSDDGKFDNTANIDRTLANALQTAKLLGFKNYAELSLATKMADTPEQVLNFLHDL
                     250       260       270       280       290       300

                     310       320       330       340       350       360
a128-1.pep    ARRAKPYAEKDLAEVKAFARESLGLADLQPWDLGYAGEKLREAKYAFSETEVKKYFPVGK
              |||||||||||||||||||||||:||||||||||||:|||||||||||||||||||||||
m128-1        ARRAKPYAEKDLAEVKAFARESLNLADLQPWDLGYASEKLREAKYAFSETEVKKYFPVGK
                     310       320       330       340       350       360

                     370       380       390       400       410       420
a128-1.pep    VLNGLFAQIKKLYGIGFTEKTVPVWHKDVRYFELQQNGETIGGVYMDLYAREGKRGGAWM
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m128-1        VLNGLFAQIKKLYGIGFTEKTVPVWHKDVRYFELQQNGETIGGVYMDLYAREGKRGGAWM
                     370       380       390       400       410       420

                     430       440       450       460       470       480
a128-1.pep    NDYKGRRRFSDGTLQLPTAYLVCNFTPPVGGKEARLSHDEILTLFHETGHGLHHLLTQVD
              |||||||||||||||||||||||||:|||||:|||||||||| |||||||||||||||||
m128-1        NDYKGRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILILFHETGHGLHHLLTQVD
                     430       440       450       460       470       480

                     490       500       510       520       530       540
a128-1.pep    ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m128-1        ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ
                     490       500       510       520       530       540

                     550       560       570       580       590       600
a128-1.pep    RGMFLVRQMEFALFDMMIYSEDDEGRLKNWQQVLDSVRKEVAVVRPPEYNRFANSFGHIF
              |||||||||||||||||||||||||||||||||||||||:|||::|||||||| ||||||
m128-1        RGMFLVRQMEFALFDMMIYSEDDEGRLKNWQQVLDSVRKKVAVIQPPEYNRFALSFGHIF
                     550       560       570       580       590       600

                     610       620       630       640       650       660
a128-1.pep    AGGYSAGYYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGGSRSAAESFKAFRG
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m128-1        AGGYSAGYYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGGSRSAAESFKAFRG
                     610       620       630       640       650       660

                     670      679
a128-1.pep    REPSIDALLRHSGFDNAAX
              |||||||||||||||||:
m128-1        REPSIDALLRHSGFDNAVX
                     670
```

206

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3097>:

m206.seq

```
  1  ATGTTTCCCC CCGACAAAAC CCTTTTCCTC TGTCTCAGCG CACTGCTCCT
 51  CGCCTCATGC GGCACGACCT CCGGCAAACA CCGCCAACCG AAACCCAAAC
101  AGACAGTCCG GCAAATCCAA GCCGTCCGCA TCAGCCACAT CGACCGCACA
151  CAAGGCTCGC AGGAACTCAT GCTCCACAGC CTCGGACTCA TCGGCACGCC
201  CTACAAATGG GGCGGCAGCA GCACCGCAAC CGGCTTCGAT TGCAGCGGCA
251  TGATTCAATT CGTTTACAAr AACGCCCTCA ACGTCAAGCT GCCGCGCACC
301  GCCCGCGACA TGGCGGCGGC AAGCCGsAAA ATCCCCGAcA GCCGCyTCAA
351  GGCCGGCGAC CTCGTATTCT TCAACACCGG CGGCGCACAC CGCTACTCAC
401  ACGTCGGACT CTACATCGGC AACGGCGAAT TCATCCATGC CCCCAGCAGC
451  GGCAAAACCA TCAAAACCGA AAAACTCTCC ACACCGTTTT ACGCCAAAAA
501  CTACCTCGGC GCACATACTT TTTTACAGA ATGA
```

This corresponds to the amino acid sequence <SEQ ID 3098; ORF 206>:

```
m206.pep..
  1  MFPPDKTLFL CLSALLLASC GTTSGKHRQP KPKQTVRQIQ AVRISHIDRT
 51  QGSQELMLHS LGLIGTPYKW GGSSTATGFD CSGMIQFVYK NALNVKLPRT
101  ARDMAAASRK IPDSRXKAGD LVFFNTGGAH RYSHVGLYIG NGEFIHAPSS
151  GKTIKTEKLS TPFYAKNYLG AHTFFTE*
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3099>:

```
g206.seq
  1  atgtttcccc ccgacaaaac ccttttcctc tgtctcggcg cactgctcct
 51  cgcctcatgc ggcacgacct ccggcaaaca ccgccaaccg aaacccaaac
101  agacagtccg gcaaatccaa gccgtccgca tcagccacat cggccgcaca
151  caaggctcgc aggaactcat gctccacagc ctcggactca tcggcacgcc
201  ctacaaatgg ggcggcagca gcaccgcaac cggcttcgac tgcagcggca
251  tgattcaatt ggtttacaaa aacgccctca acgtcaagct gccgcgcacc
301  gcccgcgaca tggcggcggc aagccgcaaa atccccgaca gccgcctcaa
351  ggccggcgac atcgtattct tcaacaccgg cggcgcacac cgctactcac
401  acgtcggact ctacatcggc aacggcgaat tcatccatgc cccggcagc
451  ggcaaaacca tcaaaaccga aaaactctcc acaccgtttt acgccaaaaa
501  ctaccttgga gcgcatacgt tttttacaga atga
```

This corresponds to the amino acid sequence <SEQ ID 3100; ORF 206.ng>:

```
g206.pep
  1  MFSPDKTLFL CLGALLLASC GTTSGKHRQP KPKQTVRQIQ AVRISHIGRT
 51  QGSQELMLHS LGLIGTPYKW GGSSTATGFD CSGMIQLVYK NALNVKLPRT
101  ARDMAAASRK IPDSRLKAGD IVFFNTGGAH RYSHVGLYIG NGEFIHAPGS
151  GKTIKTEKLS TPFYAKNYLG AHTFFTE*
```

ORF 206 shows 96.0% identity over a 177 aa overlap with a predicted ORF (ORF 206.ng) from *N. gonorrhoeae:*

```
m206/g206

                    10        20        30        40        50        60
m206.pep    MFPPDKTLFLCLSALLLASCGTTSGKHRQPKPKQTVRQIQAVRISHIDRTQGSQELMLHS
            || |||||||||:|||||||||||||||||||||||||||||||||||| ||||||||||
g206        MFSPDKTLFLCLGALLLASCGTTSGKHRQPKPKQTVRQIQAVRISHIGRTQGSQELMLHS
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m206.pep    LGLIGTPYKWGGSSTATGFDCSGMIQFVYKNALNVKLPRTARDMAAASRKIPDSRXKAGD
            |||||||||||||||||||||||||||:|||||||||||||||||||||||||||| |||
g206        LGLIGTPYKWGGSSTATGFDCSGMIQLVYKNALNVKLPRTARDMAAASRKIPDSRLKAGD
                    70        80        90       100       110       120

                   130       140       150       160       170
m206.pep    LVFFNTGGAHRYSHVGLYIGNGEFIHAPSSGKTIKTEKLSTPFYAKNYLGAHTFFTEX
            :||||||||||||||||||||||||||:||||||||||||||||||||||||||||
g206        IVFFNTGGAHRYSHVGLYIGNGEFIHAPGSGKTIKTEKLSTPFYAKNYLGAHTFFTE
```

```
        130         140         150         160         170
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3101>:

```
a206.seq
      1  ATGTTTCCCC CCGACAAAAC CCTTTTCCTC TGTCTCAGCG CACTGCTCCT
     51  CGCCTCATGC GGCACGACCT CCGGCAAACA CCGCCAACCG AAACCCAAAC
    101  AGACAGTCCG GCAAATCCAA GCCGTCCGCA TCAGCCACAT CGACCGCACA
    151  CAAGGCTCGC AGGAACTCAT GCTCCACAGC CTCGGACTCA TCGGCACGCC
    201  CTACAAATGG GGCGGCAGCA GCACCGCAAC CGGCTTCGAT TGCAGCGGCA
    251  TGATTCAATT CGTTTACAAA AACGCCCTCA ACGTCAAGCT GCCGCGCACC
    301  GCCCGCGACA TGGCGGCGGC AAGCCGCAAA ATCCCCGACA GCCGCCTTAA
    351  GGCCGGCGAC CTCGTATTCT TCAACACCGG CGGCGCACAC CGCTACTCAC
    401  ACGTCGGACT CTATATCGGC AACGGCGAAT TCATCCATGC CCCCAGCAGC
    451  GGCAAAACCA TCAAAACCGA AAAACTCTCC ACACCGTTTT ACGCCAAAAA
    501  CTACCTCGGC GCACATACTT TCTTTACAGA ATGA
```

This corresponds to the amino acid sequence <SEQ ID 3102; ORF 206.a>:

```
a206.pep
      1  MFPPDKTLFL CLSALLLASC GTTSGKHRQP KPKQTVRQIQ AVRISHIDRT
     51  QGSQELMLHS LGLIGTPYKW GGSSTATGFD CSGMIQFVYK NALNVKLPRT
    101  ARDMAAASRK IPDSRLKAGD LVFFNTGGAH RYSHVGLYIG NGEFIHAPSS
    151  GKTIKTEKLS TPFYAKNYLG AHTFFTE*
```

m206/a206 ORFs 206 and 206.a showed a 99.4% identity in 177 aa overlap

```
                    10        20        30        40        50        60
m206.pep    MFPPDKTLFLCLSALLLASCGTTSGKHRQPKPKQTVRQIQAVRISHIDRTQGSQELMLHS
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a206        MFPPDKTLFLCLSALLLASCGTTSGKHRQPKPKQTVRQIQAVRISHIDRTQGSQELMLHS
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m206.pep    LGLIGTPYKWGGSSTATGFDCSGMIQFVYKNALNVKLPRTARDMAAASRKIPDSRXKAGD
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
a206        LGLIGTPYKWGGSSTATGFDCSGMIQFVYKNALNVKLPRTARDMAAASRKIPDSRLKAGD
                    70        80        90       100       110       120

                   130       140       150       160       170
m206.pep    LVFFNTGGAHRYSHVGLYIGNGEFIHAPSSGKTIKTEKLSTPFYAKNYLGAHTFFTEX
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a206        LVFFNTGGAHRYSHVGLYIGNGEFIHAPSSGKTIKTEKLSTPFYAKNYLGAHTFFTEX
                   130       140       150       160       170
```

287

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3103>:

```
m287.seq
      1  ATGTTTAAAC GCAGCGTAAT CGCAATGGCT TGTATTTTTG CCCTTTCAGC
     51  CTGCGGGGGC GGCGGTGGCG GATCGCCCGA TGTCAAGTCG GCGGACACGC
    101  TGTCAAAACC TGCCGCCCCT GTTGTTTCTG AAAAAGAGAC AGAGGCAAAG
    151  GAAGATGCGC CACAGGCAGG TTCTCAAGGA CAGGGCGCGC CATCCGCACA
    201  AGGCAGTCAA GATATGGCGG CGGTTTCGGA AGAAAATACA GGCAATGGCG
    251  GTGCGGTAAC AGCGGATAAT CCCAAAAATG AAGACGAGGT GGCACAAAAT
    301  GATATGCCGC AAAATGCCGC CGGTACAGAT AGTTCGACAC CGAATCACAC
    351  CCCGGATCCG AATATGCTTG CCGGAAATAT GGAAAATCAA GCAACGGATG
    401  CCGGGGAATC GTCTCAGCCG GCAAACCAAC CGGATATGGC AAATGCGGCG
    451  GACGGAATGC AGGGGGACGA TCCGTCGGCA GGCGGGCAAA ATGCCGGCAA
    501  TACGGCTGCC CAAGGTGCAA ATCAAGCCGG AAACAATCAA GCCGCCGGTT
    551  CTTCAGATCC CATCCCCGCG TCAAACCCTG CACCTGCGAA TGGCGGTAGC
```

```
 601  AATTTTGGAA GGGTTGATTT GGCTAATGGC GTTTTGATTG ACGGGCCGTC
 651  GCAAAATATA ACGTTGACCC ACTGTAAAGG CGATTCTTGT AGTGGCAATA
 701  ATTTCTTGGA TGAAGAAGTA CAGCTAAAAT CAGAATTTGA AAAATTAAGT
 751  GATGCAGACA AAATAAGTAA TTACAAGAAA GATGGGAAGA ATGATAAATT
 801  TGTCGGTTTG GTTGCCGATA GTGTGCAGAT GAAGGGAATC AATCAATATA
 851  TTATCTTTTA TAAACCTAAA CCCACTTCAT TTGCGCGATT TAGGCGTTCT
 901  GCACGGTCGA GGCGGTCGCT TCCGGCCGAG ATGCCGCTGA TTCCCGTCAA
 951  TCAGGCGGAT ACGCTGATTG TCGATGGGGA AGCGGTCAGC CTGACGGGGC
1001  ATTCCGGCAA TATCTTCGCG CCCGAAGGGA ATTACCGGTA TCTGACTTAC
1051  GGGGCGGAAA AATTGCCCGG CGGATCGTAT GCCCTTCGTG TTCAAGGCGA
1101  ACCGGCAAAA GGCGAAATGC TTGCGGGCGC GGCCGTGTAC AACGGCGAAG
1151  TACTGCATTT CCATACGGAA AACGGCCGTC CGTACCCGAC CAGGGGCAGG
1201  TTTGCCGCAA AAGTCGATTT CGGCAGCAAA TCTGTGGACG GCATTATCGA
1251  CAGCGGCGAT GATTTGCATA TGGGTACGCA AAAATTCAAA GCCGCCATCG
1301  ATGGAAACGG CTTTAAGGGG ACTTGGACGG AAAATGGCAG CGGGGATGTT
1351  TCCGGAAAGT TTTACGGCCC GGCCGGCGAG GAAGTGGCGG GAAAATACAG
1401  CTATCGCCCG ACAGATGCGG AAAAGGGCGG ATTCGGCGTG TTTGCCGGCA
1451  AAAAAGAGCA GGATTGA
```

This corresponds to the amino acid sequence <SEQ ID 3104; ORF 287>:

```
m287.pep
   1  MFKRSVIAMA CIFALSACGG GGGGSPDVKS ADTLSKPAAP VVSEKETEAK
  51  EDAPQAGSQG QGAPSAQGSQ DMAAVSEENT GNGGAVTADN PKNEDEVAQN
 101  DMPQNAAGTD SSTPNHTPDP NMLAGNMENQ ATDAGESSQP ANQPDMANAA
 151  DGMQGDDPSA GGQNAGNTAA QGANQAGNNQ AAGSSDPIPA SNPAPANGGS
 201  NFGRVDLANG VLIDGPSQNI TLTHCKGDSC SGNNFLDEEV QLKSEFEKLS
 251  DADKISNYKK DGKNDKFVGL VADSVQMKGI NQYIIFYKPK PTSFARFRRS
 301  ARSRRSLPAE MPLIPVNQAD TLIVDGEAVS LTGHSGNIFA PEGNYRYLTY
 351  GAEKLPGGSY ALRVQGEPAK GEMLAGAAVY NGEVLHFHTE NGRPYPTRGR
 401  FAAKVDFGSK SVDGIIDSGD DLHMGTQKFK AAIDGNGFKG TWTENGSGDV
 451  SGKFYGPAGE EVAGKYSYRP TDAEKGGFGV FAGKKEQD*
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3105>:

```
g287.seq
   1  atgtttaaac gcagtgtgat tgcaatggct tgtatttttc ccctttcagc
  51  ctgtgggggc ggcggtggcg gatcgcccga tgtcaagtcg gcggacacgc
 101  cgtcaaaacc ggccgccccc gttgttgctg aaaatgccgg ggaaggggtg
 151  ctgccgaaag aaaagaaaga tgaggaggca gcgggcggtg cgccgcaagc
 201  cgatacgcag gacgcaaccg ccggagaagg cagccaagat atggcggcag
 251  tttcggcaga aaatacaggc aatggcggtg cggcaacaac ggacaacccc
 301  aaaaatgaag acgcgggggc gcaaaatgat atgccgcaaa atgccgccga
 351  atccgcaaat caaacaggga acaaccaacc cgccggttct tcagattccg
 401  cccccgcgtc aaaccctgcc cctgcgaatg gcggtagcga tttggaagg
 451  acgaacgtgg gcaattctgt tgtgattgac ggaccgtcgc aaaatataac
 501  gttgacccac tgtaaaggcg attcttgtaa tggtgataat ttattggatg
 551  aagaagcacc gtcaaaatca gaatttgaaa aattaagtga tgaagaaaaa
 601  attaagcgat ataaaaaaga cgagcaacgg gagaattttg tcggtttggt
 651  tgctgacagg gtaaaaaagg atggaactaa caaatatatc atcttctata
 701  cggacaaacc acctactcgt tctgcacggt cgaggaggtc gcttccggcc
 751  gagattccgc tgattcccgt caatcaggcc gatacgctga ttgtggatgg
 801  ggaagcggtc agcctgacgg ggcattccgg caatatcttc gcgcccgaag
 851  ggaattaccg gtatctgact tacggggcgg aaaaattgcc cggcggatcg
 901  tatgccctcc gtgtgcaagg cgaaccggca aaaggcgaaa tgcttgttgg
 951  cacggccgtg tacaacggcg aagtgctgca tttccatatg gaaaacggcc
1001  gtccgtaccc gtccggaggc aggtttgccg caaaagtcga tttcggcagc
1051  aaatctgtgg acggcattat cgacagcggc gatgatttgc atatgggtac
1101  gcaaaaattc aaagccgcca tcgatggaaa cggctttaag gggacttgga
1151  cggaaaatgg cggcggggat gtttccggaa ggttttacgg cccggccggc
1201  gaggaagtgg cgggaaaata cagctatcgc ccgacagatg ctgaaaaggg
1251  cggattcggc gtgtttgccg gcaaaaaaga tcgggattga
```

This corresponds to the amino acid sequence <SEQ ID 3106; ORF 287.ng>:

```
g287.pep
   1  MFKRSVIAMA CIFPLSACGG GGGGSPDVKS ADTPSKPAAP VVAENAGEGV
```

```
 51  LPKEKKDEEA AGGAPQADTQ DATAGEGSQD MAAVSAENTG NGGAATTDNP
101  KNEDAGAQND MPQNAAESAN QTGNNQPAGS SDSAPASNPA PANGGSDFGR
151  TNVGNSVVID GPSQNITLTH CKGDSCNGDN LLDEEAPSKS EFEKLSDEEK
201  IKRYKKDEQR ENFVGLVADR VKKDGTNKYI IFYTDKPPTR SARSRRSLPA
251  EIPLIPVNQA DTLIVDGEAV SLTGHSGNIF APEGNYRYLT YGAEKLPGGS
301  YALRVQGEPA KGEMLVGTAV YNGEVLHFHM ENGRPYPSGG RFAAKVDFGS
351  KSVDGIIDSG DDLHMGTQKF KAAIDGNGFK GTWTENGGGD VSGRFYGPAG
401  EEVAGKYSYR PTDAEKGGFG VFAGKKDRD*
```

m287/g287 ORFs 287 and 287.ng showed a 70.1% identity in 499 aa overlap

```
                   10        20        30        40                 49
m287.pep   MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSKPAAPVVSE-----------KETEA
           ||||||||||||| |||||||||||||||||||| |||||||:|           |: ||
g287       MFKRSVIAMACIFPLSACGGGGGGSPDVKSADTPSKPAAPVVAENAGEGVLPKEKKDEEA
                   10        20        30        40        50        60

          50        60        70        80        90       100       109
m287.pep   KEDAPQAGSQGQGAPSAQGSQDMAAVSEENTGNGGAVTADNPKNEDEVAQNDMPQNAAGT
             |||| :|     | :::||||||||| ||||||||:|:||||||| ||||||||||
g287       AGGAPQADTQD--ATAGEGSQDMAAVSAENTGNGGAATTDNPKNEDAGAQNDMPQNAA--
                   70        80        90       100       110

         110       120       130       140       150       160       169
m287.pep   DSSTPNHTPDPNMLAGNMENQATDAGESSQPANQPDMANAADGMQGDDPSAGGQNAGNTA

g287       ------------------------------------------------------------

         170       180       190       200       210       220       229
m287.pep   AQGANQAGNNQAAGSSDPIPASNPAPANGGSNFGRVDLANGVLIDGPSQNITLTHCKGDS
            ::|||:|||| |||||  ||||||||||:|:|::::|:|:||||||||||||||||||
g287       -ESANQTGNNQPAGSSDSAPASNPAPANGGSDFGRTNVGNSVVIDGPSQNITLTHCKGDS
             120       130       140       150       160       170

         230       240       250       260       270       280       289
m287.pep   CSGNNFLDEEVQLKSEFEKLSDADKISNYKKDGKNDKFVGLVADSVQMKGINQYIIFYKP
           |:|:|:||||:  ||||||||| :||: |||| : ::||||||| |:  | |:|||||
g287       CNGDNLLDEEAPSKSEFEKLSDEEKIKRYKKDEQRENFVGLVADRVKKDGTNKYIIFYTD
             180       190       200       210       220       230

         290       300       310       320       330       340       349
m287.pep   KPTSFARFRRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNYRYLT
           || :      |||||||||||:||||||||||||||||||||||||||||||||||||||
g287       KPPT-----RSARSRRSLPAEIPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNYRYLT
                  240       250       260       270       280       290

         350       360       370       380       390       400       409
m287.pep   YGAEKLPGGSYALRVQGEPAKGEMLAGAAVYNGEVLHFHTENGRPYPTRGRFAAKVDFGS
           |||||||||||||||||||||||||:|:||||||||||| |||||||: |||||||||||
g287       YGAEKLPGGSYALRVQGEPAKGEMLVGTAVYNGEVLHFHMENGRPYPSGGRFAAKVDFGS
                  300       310       320       330       340       350

         410       420       430       440       450       460       469
m287.pep   KSVDGIIDSGDDLHMGTQKFKAAIDGNGFKGTWTENGSGDVSGKFYGPAGEEVAGKYSYR
           |||||||||||||||||||||||||||||||||||||:|||||:||||||||||||||||
g287       KSVDGIIDSGDDLHMGTQKFKAAIDGNGFKGTWTENGGGDVSGRFYGPAGEEVAGKYSYR
                  360       370       380       390       400       410

         470       480       489
m287.pep   PTDAEKGGFGVFAGKKEQDX
```

```
                |||||||||||||||::||
g287            PTDAEKGGFGVFAGKKDRDX
                        420       430
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3107>:

```
a287.seq
      1  ATGTTTAAAC GCAGTGTGAT TGCAATGGCT TGTATTGTTG CCCTTTCAGC
     51  CTGTGGGGGC GGCGGTGGCG GATCGCCCGA TGTTAAGTCG GCGGACACGC
    101  TGTCAAAACC TGCCGCCCCT GTTGTTACTG AAGATGTCGG GGAAGAGGTG
    151  CTGCCGAAAG AAAAGAAAGA TGAGGAGGCG GTGAGTGGTG CGCCGCAAGC
    201  CGATACGCAG GACGCAACCG CCGGAAAAGG CGGTCAAGAT ATGGCGGCAG
    251  TTTCGGCAGA AAATACAGGC AATGGCGGTG CGGCAACAAC GGATAATCCC
    301  GAAAATAAAG ACGAGGGACC GCAAAATGAT ATGCCGCAAA ATGCCGCCGA
    351  TACAGATAGT TCGACACCGA ATCACACCCC TGCACCGAAT ATGCCAACCA
    401  GAGATATGGG AAACCAAGCA CCGGATGCCG GGGAATCGGC ACAACCGGCA
    451  AACCAACCGG ATATGGCAAA TGCGGCGGAC GGAATGCAGG GGGACGATCC
    501  GTCGGCAGGG GAAAATGCCG GCAATACGGC AGATCAAGCT GCAAATCAAG
    551  CTGAAAACAA TCAAGTCGGC GGCTCTCAAA ATCCTGCCTC TTCAACCAAT
    601  CCTAACGCCA CGAATGGCGG CAGCGATTTT GGAAGGATAA ATGTAGCTAA
    651  TGGCATCAAG CTTGACAGCG GTTCGGAAAA TGTAACGTTG ACACATTGTA
    701  AAGACAAAGT ATGCGATAGA GATTTCTTAG ATGAAGAAGC ACCACCAAAA
    751  TCAGAATTTG AAAAATTAAG TGATGAAGAA AAAATTAATA AATATAAAAA
    801  AGACGAGCAA CGAGAGAATT TTGTCGGTTT GGTTGCTGAC AGGGTAGAAA
    851  AGAATGGAAC TAACAAATAT GTCATCATTT ATAAAGACAA GTCCGCTTCA
    901  TCTTCATCTG CGCGATTCAG GCGTTCTGCA CGGTCGAGGC GGTCGCTTCC
    951  GGCCGAGATG CCGCTGATTC CCGTCAATCA GGCGGATACG CTGATTGTCG
   1001  ATGGGGAAGC GGTCAGCCTG ACGGGGCATT CCGGCAATAT CTTCGCGCCC
   1051  GAAGGGAATT ACCGGTATCT GACTTACGGG GCGGAAAAAT TGTCCGGCGG
   1101  ATCGTATGCC CTCAGTGTGC AAGGCGAACC GGCAAAAGGC GAAATGCTTG
   1151  CGGGCACGGC CGTGTACAAC GGCGAAGTGC TGCATTTCCA TATGGAAAAC
   1201  GGCCGTCCGT CCCCGTCCGG AGGCAGGTTT GCCGCAAAAG TCGATTTCGG
   1251  CAGCAAATCT GTGGACGGCA TTATCGACAG CGGCGATGAT TTGCATATGG
   1301  GTACGCAAAA ATTCAAAGCC GTTATCGATG GAAACGGCTT TAAGGGGACT
   1351  TGGACGGAAA ATGGCGGCGG GGATGTTTCC GGAAGGTTTT ACGGCCCGGC
   1401  CGGCGAAGAA GTGGCGGGAA AATACAGCTA TCGCCCGACA GATGCGGAAA
   1451  AGGGCGGATT CGGCGTGTTT GCCGGCAAAA AAGAGCAGGA TTGA
```

This corresponds to the amino acid sequence <SEQ ID 3108; ORF 287.a>:

```
a287.pep
      1  MFKRSVIAMA CIVALSACGG GGGGSPDVKS ADTLSKPAAP VVTEDVGEEV
     51  LPKEKKDEEA VSGAPQADTQ DATAGKGGQD MAAVSAENTG NGGAATTDNP
    101  ENKDEGPQND MPQNAADTDS STPNHTPAPN MPTRDMGNQA PDAGESAQPA
    151  NQPDMANAAD GMQGDDPSAG ENAGNTADQA ANQAENNQVG GSQNPASSTN
    201  PNATNGGSDF GRINVANGIK LDSGSENVTL THCKDKVCDR DFLDEEAPPK
    251  SEFEKLSDEE KINKYKKDEQ RENFVGLVAD RVEKNGTNKY VIIYKDKSAS
    301  SSSARFRRSA RSRRSLPAEM PLIPVNQADT LIVDGEAVSL TGHSGNIFAP
    351  EGNYRYLTYG AEKLSGGSYA LSVQGEPAKG EMLAGTAVYN GEVLHFHMEN
    401  GRPSPSGGRF AAKVDFGSKS VDGIIDSGDD LHMGTQKFKA VIDGNGFKGT
    451  WTENGGGDVS GRFYGPAGEE VAGKYSYRPT DAEKGGFGVF AGKKEQD*
```

```
m287/a287     ORFs 287 and 287.a showed a 77.2% identity in 501 aa overlap

                        10        20        30        40                 49
m287.pep        MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSKPAAPVVSE-----------KETEA
                |||||||||||| |||||||||||||||||||||||||||||:|           |: ||
a287            MFKRSVIAMACIVALSACGGGGGGSPDVKSADTLSKPAAPVVTEDVGEEVLPKEKKDEEA
                        10        20        30        40        50        60

              50        60        70        80        90       100       109
m287.pep        KEDAPQAGSQGQGAPSAQGSQDMAAVSEENTGNGGAVTADNPKNEDEVAQNDMPQNAAGT
                   |||| :|   | :::|:||||||| ||||||||:|:|||:|:||  ||||||||| |
a287            VSGAPQADTQ--DATAGKGGQDMAAVSAENTGNGGAATTDNPENKDEGPQNDMPQNAADT
                        70        80        90       100       110
```

```
               110       120       130       140       150       160      169
m287.pep  DSSTPNHTPDPNMLAGNMENQATDAGESSQPANQPDMANAADGMQGDDPSAGGQNAGNTA
          ||||||||| ||| : :| ||| |||||:||||||||||||||||||||||| :||||||
a287      DSSTPNHTPAPNMPTRDMGNQAPDAGESAQPANQPDMANAADGMQGDDPSAG-ENAGNTA
                120       130       140       150       160       170

               170       180       190       200       210       220      229
m287.pep  AQGANQAGNNQAAGSSDPIPASNPAPANGGSNFGRVDLANGVLIDGPSQNITLTHCKGDS
           |:|||| |||::||::|  ::||  :||||:|||:::|||: :|: |:|:||||||
a287      DQAANQAENNQVGGSQNPASSTNPNATNGGSDFGRINVANGIKLDSGSENVTLTHCKDKV
          180       190       200       210       220       230

               230       240       250       260       270       280      289
m287.pep  CSGNNFLDEEVQLKSEFEKLSDADKISNYKKDGKNDKFVGLVADSVQMKGINQYIIFYKP
          |:  :|||||:  ||||||||| :||::|||| : ::||||||| |: :| |:|:|:||
a287      CD-RDFLDEEAPPKSEFEKLSDEEKINKYKKDEQRENFVGLVADRVEKNGTNKYVIIYKD
           240       250       260       270       280       290

               290       300       310       320       330       340
m287.pep  KP--TSFARFRRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNYRY
          |   :| |||||||||||||.||||||||||||||||||||||||||||||||||||||||
a287      KSASSSSARFRRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNYRY
           300       310       320       330       340       350

          350       360       370       380       390       400
m287.pep  LTYGAEKLPGGSYALRVQGEPAKGEMLAGAAVYNGEVLHFHTENGRPYPTRGRFAAKVDF
          |||||||| |||||| |||||||||||||:||||||||||| ||||| |: |||||||||
a287      LTYGAEKLSGGSYALSVQGEPAKGEMLAGTAVYNGEVLHFHMENGRPSPSGGRFAAKVDF
           360       370       380       390       400       410

          410       420       430       440       450       460
m287.pep  GSKSVDGIIDSGDDLHMGTQKFKAAIDGNGFKGTWTENGSGDVSGKFYGPAGEEVAGKYS
          ||||||||||||||||||||||||:||||||||||||||:|||||:||||||||||||||
a287      GSKSVDGIIDSGDDLHMGTQKFKAVIDGNGFKGTWTENGGGDVSGRFYGPAGEEVAGKYS
           420       430       440       450       460       470

          470       480       489
m287.pep  YRPTDAEKGGFGVFAGKKEQDX
          ||||||||||||||||||||||
a287      YRPTDAEKGGFGVFAGKKEQDX
           480       490
```

406

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3109>:

m406.seq

```
  1  ATGCAAGCAC GGCTGCTGAT ACCTATTCTT TTTTCAGTTT TTATTTTATC
 51  CGCCTGCGGG ACACTGACAG GTATTCCATC GCATGGCGGA GGTAAACGCT
101  TTGCGGTCGA ACAAGAACTT GTGGCCGCTT CTGCCAGAGC TGCCGTTAAA
151  GACATGGATT TACAGGCATT ACACGGACGA AAAGTTGCAT TGTACATTGC
201  CACTATGGGC GACCAAGGTT CAGGCAGTTT GACAGGGGGT CGCTACTCCA
251  TTGATGCACT GATTCGTGGC GAATACATAA ACAGCCCTGC CGTCCGTACC
301  GATTACACCT ATCCACGTTA CGAAACCACC GCTGAAACAA CATCAGGCGG
351  TTTGACAGGT TTAACCACTT CTTTATCTAC ACTTAATGCC CCTGCACTCT
401  CTCGCACCCA ATCAGACGGT AGCGGAAGTA AAAGCAGTCT GGGCTTAAAT
451  ATTGGCGGGA TGGGGGATTA TCGAAATGAA ACCTTGACGA CTAACCCGCG
501  CGACACTGCC TTTCTTTCCC ACTTGGTACA GACCGTATTT TTCCTGCGCG
551  GCATAGACGT TGTTTCTCCT GCCAATGCCG ATACAGATGT GTTTATTAAC
601  ATCGACGTAT TCGGAACGAT ACGCAACAGA ACCGAAATGC ACCTATACAA
651  TGCCGAAACA CTGAAAGCCC AAACAAAACT GGAATATTTC GCAGTAGACA
```

```
701  GAACCAATAA AAAATTGCTC ATCAAACCAA AAACCAATGC GTTTGAAGCT
751  GCCTATAAAG AAAATTACGC ATTGTGGATG GGGCCGTATA AAGTAAGCAA
801  AGGAATTAAA CCGACGGAAG GATTAATGGT CGATTTCTCC GATATCCGAC
851  CATACGGCAA TCATACGGGT AACTCCGCCC CATCCGTAGA GGCTGATAAC
901  AGTCATGAGG GGTATGGATA CAGCGATGAA GTAGTGCGAC AACATAGACA
951  AGGACAACCT TGA
```

This corresponds to the amino acid sequence <SEQ ID 3110; ORF 406>:

m406.pep

```
  1  MQARLLIPIL FSVFILSACG TLTGIPSHGG GKRFAVEQEL VAASARAAVK
 51  DMDLQALHGR KVALYIATMG DQGSGSLTGG RYSIDALIRG EYINSPAVRT
101  DYTYPRYETT AETTSGGLTG LTTSLSTLNA PALSRTQSDG SGSKSSLGLN
151  IGGMGDYRNE TLTTNPRDTA FLSHLVQTVF FLRGIDVVSP ANADTDVFIN
201  IDVFGTIRNR TEMHLYNAET LKAQTKLEYF AVDRTNKKLL IKPKTNAFEA
251  AYKENYALWM GPYKVSKGIK PTEGLMVDFS DIRPYGNHTG NSAPSVEADN
301  SHEGYGYSDE VVRQHRQGQP *
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3111>:

g406.seq

```
  1  ATGCGGGCAC GGCTGCTGAT ACCTATTCTT TTTTCAGTTT TTATTTTATC
 51  CGCCTGCGGG ACACTGACAG GTATTCCATC GCATGGCGGA GGCAAACGCT
101  TCGCGGTCGA ACAAGAACTT GTGGCCGCTT CTGCCAGAGC TGCCGTTAAA
151  GACATGGATT TACAGGCATT ACACGGACGA AAAGTTGCAT TGTACATTGC
201  AACTATGGGC GACCAAGGTT CAGGCAGTTT GACAGGGGGT CGCTACTCCA
251  TTGATGCACT GATTCGCGGC GAATACATAA ACAGCCCTGC CGTCCGCACC
301  GATTACACCT ATCCGCGTTA CGAAACCACC GCTGAAACAA CATCAGGCGG
351  TTTGACGGGT TTAACCACTT CTTTATCTAC ACTTAATGCC CCTGCACTCT
401  CGCGCACCCA ATCAGACGGT AGCGGAAGTA GGAGCAGTCT GGGCTTAAAT
451  ATTGGCGGGA TGGGGGATTA TCGAAATGAA ACCTTGACGA CCAACCCGCG
501  CGACACTGCC TTTCTTTCCC ACTTGGTGCA GACCGTATTT TTCCTGCGCG
551  GCATAGACGT TGTTTCTCCT GCCAATGCCG ATACAGATGT GTTTATTAAC
601  ATCGACGTAT TCGGAACGAT ACGCAACAGA ACCGAAATGC ACCTATACAA
651  TGCCGAAACA CTGAAAGCCC AAACAAAACT GGAATATTTC GCAGTAGACA
701  GAACCAATAA AAAATTGCTC ATCAAACCCA AAACCAATGC GTTTGAAGCT
751  GCCTATAAAG AAAATTACGC ATTGTGGATG GGGCCGTATA AAGTAAGCAA
801  AGGAATCAAA CCGACGGAAG GATTGATGGT CGATTTCTCC GATATCCAAC
851  CATACGGCAA TCATACGGGT AACTCCGCCC CATCCGTAGA GGCTGATAAC
901  AGTCATGAGG GGTATGGATA CAGCGATGAA GCAGTGCGAC AACATAGACA
951  AGGGCAACCT TGA
```

This corresponds to the amino acid sequence <SEQ ID 3112; ORF 406>:

g406.pep

```
  1  MRARLLIPIL FSVFILSACG TLTGIPSHGG GKRFAVEQEL VAASARAAVK
 51  DMDLQALHGR KVALYIATMG DQGSGSLTGG RYSIDALIRG EYINSPAVRT
101  DYTYPRYETT AETTSGGLTG LTTSLSTLNA PALSRTQSDG SGSRSSLGLN
151  IGGMGDYRNE TLTTNPRDTA FLSHLVQTVF FLRGIDVVSP ANADTDVFIN
201  IDVFGTIRNR TEMHLYNAET LKAQTKLEYF AVDRTNKKLL IKPKTNAFEA
251  AYKENYALWM GPYKVSKGIK PTEGLMVDFS DIQPYGNHTG NSAPSVEADN
301  SHEGYGYSDE AVRQHRQGQP *
```

ORF 406 shows 98.8% identity over a 320 aa overlap with a predicted ORF (ORF406.a) from *N. gonorrhoeae:*

g406/m406

```
                   10        20        30        40        50        60
g406.pep  MRARLLIPILFSVFILSACGTLTGIPSHGGGKRFAVEQELVAASARAAVKDMDLQALHGR
          |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m406      MQARLLIPILFSVFILSACGTLTGIPSHGGGKRFAVEQELVAASARAAVKDMDLQALHGR
```

```
                   10         20         30         40         50         60

                   70         80         90        100        110        120
g406.pep    KVALYIATMGDQGSGSLTGGRYSIDALIRGEYINSPAVRTDYTYPRYETTAETTSGGLTG
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m406        KVALYIATMGDQGSGSLTGGRYSIDALIRGEYINSPAVRTDYTYPRYETTAETTSGGLTG
                   70         80         90        100        110        120

                  130        140        150        160        170        180
g406.pep    LTTSLSTLNAPALSRTQSDGSGSRSSLGLNIGGMGDYRNETLTTNPRDTAFLSHLVQTVF
            ||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
m406        LTTSLSTLNAPALSRTQSDGSGSKSSLGLNIGGMGDYRNETLTTNPRDTAFLSHLVQTVF
                  130        140        150        160        170        180

                  190        200        210        220        230        240
g406.pep    FLRGIDVVSPANADTDVFINIDVFGTIRNRTEMHLYNAETLKAQTKLEYFAVDRTNKKLL
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m406        FLRGIDVVSPANADTDVFINIDVFGTIRNRTEMHLYNAETLKAQTKLEYFAVDRTNKKLL
                  190        200        210        220        230        240

                  250        260        270        280        290        300
g406.pep    IKPKTNAFEAAYKENYALWMGPYKVSKGIKPTEGLMVDFSDIQPYGNHTGNSAPSVEADN
            |||||||||||||||||||||||||||||||||||||||:|||||:||||||||||||||||
m406        IKPKTNAFEAAYKENYALWMGPYKVSKGIKPTEGLMVDFSDIRPYGNHTGNSAPSVEADN
                  250        260        270        280        290        300

                  310        320
g406.pep    SHEGYGYSDEAVRQHRQGQPX
            ||||||||||:||||||||||
m406        SHEGYGYSDEVVRQHRQGQPX
                  310        320
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3113>:

```
a406.seq
      1  ATGCAAGCAC GGCTGCTGAT ACCTATTCTT TTTTCAGTTT TTATTTTATC
     51  CGCCTGCGGG ACACTGACAG GTATTCCATC GCATGGCGGA GGTAAACGCT
    101  TCGCGGTCGA ACAAGAACTT GTGGCCGCTT CTGCCAGAGC TGCCGTTAAA
    151  GACATGGATT TACAGGCATT ACACGGACGA AAAGTTGCAT TGTACATTGC
    201  AACTATGGGC GACCAAGGTT CAGGCAGTTT GACAGGGGGT CGCTACTCCA
    251  TTGATGCACT GATTCGTGGC GAATACATAA ACAGCCCTGC CGTCCGTACC
    301  GATTACACCT ATCCACGTTA CGAAACCACC GCTGAAACAA CATCAGGCGG
    351  TTTGACAGGT TTAACCACTT CTTTATCTAC ACTTAATGCC CCTGCACTCT
    401  CGCGCACCCA ATCAGACGGT AGCGGAAGTA AAAGCAGTCT GGGCTTAAAT
    451  ATTGGCGGGA TGGGGGATTA TCGAAATGAA ACCTTGACGA CTAACCCGCG
    501  CGACACTGCC TTTCTTTCCC ACTTGGTACA GACCGTATTT TTCCTGCGCG
    551  GCATAGACGT TGTTTCTCCT GCCAATGCCG ATACGGATGT GTTTATTAAC
    601  ATCGACGTAT TCGGAACGAT ACGCAACAGA ACCGAAATGC ACCTATACAA
    651  TGCCGAAACA CTGAAAGCCC AAACAAAACT GGAATATTTC GCAGTAGACA
    701  GAACCAATAA AAAATTGCTC ATCAAACCAA AAACCAATGC GTTTGAAGCT
    751  GCCTATAAAG AAAATTACGC ATTGTGGATG GGACCGTATA AAGTAAGCAA
    801  AGGAATTAAA CCGACAGAAG GATTAATGGT CGATTTCTCC GATATCCAAC
    851  CATACGGCAA TCATATGGGT AACTCTGCCC CATCCGTAGA GGCTGATAAC
    901  AGTCATGAGG GGTATGGATA CAGCGATGAA GCAGTGCGAC GACATAGACA
    951  AGGGCAACCT TGA
```

This corresponds to the amino acid sequence <SEQ ID 3114; ORF 406.a>:

```
a406.pep
      1  MQARLLIPIL FSVFILSACG TLTGIPSHGG GKRFAVEQEL VAASARAAVK
     51  DMDLQALHGR KVALYIATMG DQGSGSLTGG RYSIDALIRG EYINSPAVRT
    101  DYTYPRYETT AETTSGGLTG LTTSLSTLNA PALSRTQSDG SGSKSSLGLN
    151  IGGMGDYRNE TLTTNPRDTA FLSHLVQTVF FLRGIDVVSP ANADTDVFIN
    201  IDVFGTIRNR TEMHLYNAET LKAQTKLEYF AVDRTNKKLL IKPKTNAFEA
```

```
251  AYKENYALWM GPYKVSKGIK PTEGLMVDFS DIQPYGNHMG NSAPSVEADN
301  SHEGYGYSDE AVRRHRQGQP *

m406/a406    ORFs 406 and 406.a showed a 98.8% identity in 320 aa overlap

                      10        20        30        40        50        60
m406.pep     MQARLLIPILFSVFILSACGTLTGIPSHGGGKRFAVEQELVAASARAAVKDMDLQALHGR
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a406         MQARLLIPILFSVFILSACGTLTGIPSHGGGKRFAVEQELVAASARAAVKDMDLQALHGR
                      10        20        30        40        50        60

                      70        80        90       100       110       120
m406.pep     KVALYIATMGDQGSGSLTGGRYSIDALIRGEYINSPAVRTDYTYPRYETTAETTSGGLTG
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a406         KVALYIATMGDQGSGSLTGGRYSIDALIRGEYINSPAVRTDYTYPRYETTAETTSGGLTG
                      70        80        90       100       110       120

                     130       140       150       160       170       180
m406.pep     LTTSLSTLNAPALSRTQSDGSGSKSSLGLNIGGMGDYRNETLTTNPRDTAFLSHLVQTVF
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a406         LTTSLSTLNAPALSRTQSDGSGSKSSLGLNIGGMGDYRNETLTTNPRDTAFLSHLVQTVF
                     130       140       150       160       170       180

                     190       200       210       220       230       240
m406.pep     FLRGIDVVSPANADTDVFINIDVFGTIRNRTEMHLYNAETLKAQTKLEYFAVDRTNKKLL
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a406         FLRGIDVVSPANADTDVFINIDVFGTIRNRTEMHLYNAETLKAQTKLEYFAVDRTNKKLL
                     190       200       210       220       230       240

                     250       260       270       280       290       300
m406.pep     IKPKTNAFEAAYKENYALWMGPYKVSKGIKPTEGLMVDFSDIRPYGNHTGNSAPSVEADN
             ||||||||||||||||||||||||||||||||||||||||||:|||||| ||||||||||
a406         IKPKTNAFEAAYKENYALWMGPYKVSKGIKPTEGLMVDFSDIQPYGNHMGNSAPSVEADN
                     250       260       270       280       290       300

                     310       320
m406.pep     SHEGYGYSDEVVRQHRQGQPX
             ||||||||||:||:|||||||
a406         SHEGYGYSDEAVRRHRQGQPX
                     310       320
```

EXAMPLE 2

Expression of ORF 919

The primer described in Table 1 for ORF 919 was used to locate and clone ORF 919. The predicted gene *919* was cloned in pET vector and expressed in *E. coli*. The product of protein expression and purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 919-His fusion protein purification. Mice were immunized with the purified 919-His and sera were used for Western blot (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Symbols: M1, molecular weight marker; PP, purified protein, TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm

that 919 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 919 are provided in Figure 10. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 919 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 3

Expression of ORF 279

The primer described in Table 1 for ORF 279 was used to locate and clone ORF 279. The predicted gene *279* was cloned in pGex vector and expressed in *E. coli*. The product of protein expression and purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 279-GST purification. Mice were immunized with the purified 279-GST and sera were used for Western blot analysis (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vescicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 279 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 279 are provided in Figure 11. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 279 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 4

Expression of ORF 576 and 576-1

The primer described in Table 1 for ORF 576 was used to locate and clone ORF 576. The predicted gene *576* was cloned in pGex vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 576-GST fusion protein purification. Mice were immunized with the purified 576-GST and sera

were used for Western blot (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vescicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B).. These experiments confirm that ORF 576 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 576 are provided in Figure 12. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 576 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 5

Expression of ORF 519 and 519-1

The primer described in Table 1 for ORF 519 was used to locate and clone ORF 519. The predicted gene *519* was cloned in pET vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 519-His fusion protein purification. Mice were immunized with the purified 519-His and sera were used for Western blot (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 519 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 519 are provided in Figure 13. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 519 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 6

Expression of ORF 121 and 121-1

The primer described in Table 1 for ORF 121 was used to locate and clone ORF 121. The predicted gene *121* was cloned in pET vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 121-His fusion protein purification. Mice were immunized with the purified 121-His and sera were used for Western blot analysis (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Results show that 121 is a surface-exposed protein. Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vescicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 121 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 121 are provided in Figure 14. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 121 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 7

Expression of ORF 128 and 128-1

The primer described in Table 1 for ORF 128 was used to locate and clone ORF 128. The predicted gene *128* was cloned in pET vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 128-His purification. Mice were immunized with the purified 128-His and sera were used for Western blot analysis (panel B), FACS analysis (panel C), bactericidal assay (panel D) and ELISA assay (panel E). Results show that 128 is a surface-exposed protein. Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 128 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 128 are provided in Figure 15. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J.*

Immunol 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 128 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 8

Expression of ORF 206

The primer described in Table 1 for ORF 206 was used to locate and clone ORF 206. The predicted gene *206* was cloned in pET vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 206-His purification. Mice were immunized with the purified 206-His and sera were used for Western blot analysis (panel B). It is worthnoting that the immunoreactive band in protein extracts from meningococcus is 38 kDa instead of 17 kDa (panel A). To gain information on the nature of this antibody staining we expressed ORF 206 in *E. coli* without the His-tag and including the predicted leader peptide. Western blot analysis on total protein extracts from *E. coli* expressing this native form of the 206 protein showed a recative band at a position of 38 kDa, as observed in meningococcus. We conclude that the 38 kDa band in panel B) is specific and that anti-206 antibodies, likely recognize a multimeric protein complex. In panel C is shown the FACS analysis, in panel D the bactericidal assay, and in panel E) the ELISA assay. Results show that 206 is a surface-exposed protein. Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 206 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 519 are provided in Figure 16. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 206 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 9

Expression of ORF 287

The primer described in Table 1 for ORF 287 was used to locate and clone ORF 287. The predicted gene *287* was cloned in pGex vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 287-GST fusion protein purification. Mice were immunized with the purified 287-GST and sera were used for FACS analysis (panel B), bactericidal assay (panel C), and ELISA assay (panel D). Results show that 287 is a surface-exposed protein. Symbols: M1, molecular weight marker. Arrow indicates the position of the main recombinant protein product (A). These experiments confirm that 287 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 287 are provided in Figure 17. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 287 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 10

Expression of ORF 406

The primer described in Table 1 for ORF 406 was used to locate and clone ORF 406. The predicted gene *406* was cloned in pET vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 406-His fusion protein purification. Mice were immunized with the purified 406-His and sera were used for Western blot analysis (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Results show that 406 is a surface-exposed protein. Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vescicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 406 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 406 are provided in Figure 18. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al.

1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 406 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 11

Table 2 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 225 among different strains.

Table 2

225 gene variability: List of used Neisseria strains

Identification number	Strains	Source / reference
Group B		
zo01_225	NG6/88	R. Moxon / Seiler *et al.*, 1996
zo02_225	BZ198	R. Moxon / Seiler *et al.*, 1996
zo03_225	NG3/88	R. Moxon / Seiler *et al.*, 1996
zo04_225	297-0	R. Moxon / Seiler *et al.*, 1996
zo05_225	1000	R. Moxon / Seiler *et al.*, 1996
zo06_225	BZ147	R. Moxon / Seiler *et al.*, 1996
zo07_225	BZ169	R. Moxon / Seiler *et al.*, 1996
zo08_225	528	R. Moxon / Seiler *et al.*, 1996
zo09_225	NGP165	R. Moxon / Seiler *et al.*, 1996
zo10_225	BZ133	R. Moxon / Seiler *et al.*, 1996
zo11_225	NGE31	R. Moxon / Seiler *et al.*, 1996
zo12_225	NGF26	R. Moxon / Seiler *et al.*, 1996
zo13_225	NGE28	R. Moxon / Seiler *et al.*, 1996
zo14_225	NGH38	R. Moxon / Seiler *et al.*, 1996
zo15_225	SWZ107	R. Moxon / Seiler *et al.*, 1996
zo16_225	NGH15	R. Moxon / Seiler *et al.*, 1996
zo17_225	NGH36	R. Moxon / Seiler *et al.*, 1996
zo18_225	BZ232	R. Moxon / Seiler *et al.*, 1996
zo19_225	BZ83	R. Moxon / Seiler *et al.*, 1996
zo20_225	44/76	R. Moxon / Seiler *et al.*, 1996
zo21_225	MC58	R. Moxon
zo96_225	2996	Our collection
Group A		
zo22_225	205900	R. Moxon
zo23_225	F6124	R. Moxon
z2491	Z2491	R. Moxon / Maiden *et al.*, 1998
Group C		
zo24_225	90/18311	R. Moxon
zo25_225	93/4286	R. Moxon

Others

zo26_225 A22 (group W) R. Moxon / Maiden *et al.*, 1998
zo27_225 E26 (group X) R. Moxon / Maiden *et al.*, 1998
zo28_225 860800 (group Y) R. Moxon / Maiden *et al.*, 1998
zo29_225 E32 (group Z) R. Moxon / Maiden *et al.*, 1998

Gonococcus

zo32_225 Ng F62 R. Moxon / Maiden *et al.*, 1998
zo33_225 Ng SN4 R. Moxon

fa1090 FA1090 R. Moxon

References:

Seiler A. *et al.*, Mol. Microbiol., 1996, 19(4):841-856.
Maiden et al., Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.

The amino acid sequences for each listed strain are as follows:

```
>FA1090 <SEQ ID 3115>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPVNRAPARRAG
NADELIGSAMGLNEQPVLPVNRAPARRAGNADELIGSAMGLLGIAYRYGGTSVSTGFDCS
GFMQHIFKRAMGINLPRTSAEQARMGAPVARSELQPGDMVFFRTLGGSRISHVGLYIGNN
RFIHAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDPSRFLN*

Z2491 <SEQ ID 3116>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRVPARRAGNA
DELIGNAMGLNEQPVLPVNRAPARRAGNADELIGNAMGLLGIAYRYGGTSISTGFDCSGF
MQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNNRF
IHAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDPSRFLN*

Z001_225 <SEQ ID 3117>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

Z002_225 <SEQ ID 3118>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

Z003_225 <SEQ ID 3119>
MDSFFKPAVWAVLWLMFAVRLALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*
```

Z004_225 <SEQ ID 3120>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

Z005_225 <SEQ ID 3121>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGSAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

Z006_225 <SEQ ID 3122>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

Z007_225 <SEQ ID 3123>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

Z008_225 <SEQ ID 3124>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGSAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

Z009_225 <SEQ ID 3125>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

Z010_225 <SEQ ID 3126>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

Z011_225 <SEQ ID 3127>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLNEQPVLPVNRAPARRAGNADELIGNAMGLLGIAYRYGGTSVSTGFDCSGF
MQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNNRF
IHAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDPSRFLN*

Z012_225 <SEQ ID 3128>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

Z013_225 <SEQ ID 3129>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPVNRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFIQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

Z014_225 <SEQ ID 3130>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGFNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

Z015_225 <SEQ ID 3131>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLLGIAYRYGGTSVSTGFDCS
GFMQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNN
RFIHAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDPSRFLN*

Z016_225 <SEQ ID 3132>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

Z017_225 <SEQ ID 3133>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

Z018_225 <SEQ ID 3134>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

Z019_225 <SEQ ID 3135>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

Z020_225 <SEQ ID 3136>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPINRAPARRAGNADELIGSAMGLNEQPVLPVNRVPARRAGNA
DELIGNAMGLNEQPVLPVNRAPARRAGNADELIGNAMGLLGIAYRYGGTSVSTGFDCSGF
MQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNNRF
IHAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDPSRFLN*

Z021_225 <SEQ ID 3137>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

Z022_225 <SEQ ID 3138>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR

SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

ZO23_225 <SEQ ID 3139>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

ZO24_225 <SEQ ID 3140>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

ZO25_225 <SEQ ID 3141>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

ZO26_225 <SEQ ID 3142>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

ZO27_225 <SEQ ID 3143>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

ZO28_225 <SEQ ID 3144>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

ZO29_225 <SEQ ID 3145>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

ZO32_225 <SEQ ID 3146>
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPVNRAPARRAG
NADELIGSAMGLNEQPVLPVNRAPARRAGNADELIGSAMGLLGIAYRYGGTSVSTGFDCS
GFMQHIFKRAMGINLPRTSAEQARMGAPVARSELQPGDMVFFRTLGGSRISHVGLYIGNN
RFIHAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDPSRFLN*

ZO33_225 <SEQ ID 3147>
MDSFFKPAVWAVLWLMFAVRSALADELTNLLSSREQILRQFAEDEQPVLPVNRAPARRAG
NADELIGSAMGLNEQPVLPVNRAPARRAGNADELIGSAMGLLGIAYRYGGTSVSTGFDCS
GFMQHIFKRAMGINLPRTSAEQARMGAPVARSELQPGDMVFFRTLGGSRISHVGLYIGNN
RFIHAPRTGKNIEITSLSHKYWSGKYAFARRIKKNDPSRFLN*

ZO96_225 <SEQ ID 3148>

```
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*
```

Figure 19 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 225, further confirming its utility as an antigen for both vaccines and diagnostics.

EXAMPLE 12

Table 3 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 235 among different strains.

Table 3

235 gene variability: List of used Neisseria strains		
Identification number	**Strains**	**Reference**
	Group B	
gnmzq01	NG6/88	Seiler *et al.*, 1996
gnmzq02	BZ198	Seiler *et al.*, 1996
gnmzq03	NG3/88	Seiler *et al.*, 1996
gnmzq04	1000	Seiler *et al.*, 1996
gnmzq05	1000	Seiler *et al.*, 1996
gnmzq07	BZ169	Seiler *et al.*, 1996
gnmzq08	528	Seiler *et al.*, 1996
gnmzq09	NGP165	Seiler *et al.*, 1996
gnmzq10	BZ133	Seiler *et al.*, 1996
gnmzq11	NGE31	Seiler *et al.*, 1996
gnmzq13	NGE28	Seiler *et al.*, 1996
gnmzq14	NGH38	Seiler *et al.*, 1996
gnmzq15	SWZ107	Seiler *et al.*, 1996
gnmzq16	NGH15	Seiler *et al.*, 1996
gnmzq17	NGH36	Seiler *et al.*, 1996
gnmzq18	BZ232	Seiler *et al.*, 1996
gnmzq19	BZ83	Seiler *et al.*, 1996
gnmzq21	MC58	Virji *et al.*, 1992
	Group A	
gnmzq22	205900	Our collection

gnmzq23	F6124	Our collection
z2491	Z2491	Maiden *et al.*, 1998
	Group C	
gnmzq24	90/18311	Our collection
gnmzq25	93/4286	Our collection
	Others	
gnmzq26	A22 (group W)	Maiden *et al.*, 1998
gnmzq27	E26 (group X)	Maiden *et al.*, 1998
gnmzq28	860800 (group Y)	Maiden *et al.*, 1998
gnmzq29	E32 (group Z)	Maiden *et al.*, 1998
gnmzq31	*N. lactamica*	Our collection
	Gonococcus	
gnmzq32	Ng F62	Maiden *et al.*, 1998
gnmzq33	Ng SN4	Our collection
fa1090	FA1090	Dempsey *et al.* 1991

References:

Seiler A. *et al.*, Mol. Microbiol., 1996, 19(4):841-856.
Maiden R. *et al.*, Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.
Virji M. *et al.*, Mol. Microbiol., 1992, 6:1271-1279
Dempsey J.F. *et al.*, J. Bacteriol., 1991, 173:5476-5486

The amino acid sequences for each listed strain are as follows:

```
FA1090 <SEQ ID 3149>
MKPLILGLAAVLALSACQVRKAPDLDYTSFKESKPASILVVPPLNESPDVNGTWGMLAST
AAPISEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVGAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSRNGILKGPRFVEEQPK*

GNMZQ01 <SEQ ID 3150>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANNLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ02 <SEQ ID 3151>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ03 <SEQ ID 3152>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
```

AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ04 <SEQ ID 3153>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ05 <SEQ ID 3154>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANNLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ07 <SEQ ID 3155>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ08 <SEQ ID 3156>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANNLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ09 <SEQ ID 3157>
MKPLILGLAAALVLSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGMLAST
AEPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVQPEKLHQIFGNDAVLYITITEYGTS
YQILDSVTTVSARARLVDSRNGKVLWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ10 <SEQ ID 3158>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ11 <SEQ ID 3159>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ13 <SEQ ID 3160>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ14 <SEQ ID 3161>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVGAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ15 <SEQ ID 3162>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ16 <SEQ ID 3163>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ17 <SEQ ID 3164>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ18 <SEQ ID 3165>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVGAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ19 <SEQ ID 3166>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ21 <SEQ ID 3166>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ22 <SEQ ID 3167>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ23 <SEQ ID 3168>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ24 <SEQ ID 3169>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ25 <SEQ ID 3170>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ26 <SEQ ID 3171>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGMLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVGAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ27 <SEQ ID 3172>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ28 <SEQ ID 3173>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ29 <SEQ ID 3174>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ31 <SEQ ID 3175>
MKPLILGLAAVLALSACQVQKAPDFDYTAFKESKPASILVVPPLNESPDVNGTWGMLAST
AEPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITITEYGTS
YQILDSVTTVSARARLVDSRNGKVLWSGSASIREGSNNSNSGLLGALVGAVVNQIANSLT
DRGYQVSKAAAYDLLSPYSHNGILKGPRFVEEQPK*

GNMZQ32 <SEQ ID 3176>
MKPLILGLAAVLALSACQVRKAPDLDYTSFKESKPASILVVPPLNESPDVNGTWGMLAST
AAPISEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVGAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSRNGILKGPRFVEEQPK*

GNMZQ33 <SEQ ID 3177>
MKPLILGLAAVLALSACQVRKAPDLDYTSFKESKPASILVVPPLNESPDVNGTWGMLAST
AAPISEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVGAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSRNGILKGPRFVEEQPK*

Z2491 <SEQ ID 3178>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

Figure 20 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 235, further confirming its utility as an antigen for both vaccines and diagnostics.

EXAMPLE 13

Table 4 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 287 among different strains.

Table 4

287 gene variability: List of used Neisseria strains	
Identification Strains number	**Reference**

	Group B	
287_2	BZ198	Seiler *et al.*, 1996
287_9	NGP165	Seiler *et al.*, 1996
287_14	NGH38	Seiler *et al.*, 1996
287_21	MC58	Virji *et al.*, 1992
	Group A	
z2491	Z2491	Maiden *et al.*, 1998
	Gonococcus	
fa1090	FA1090	Dempsey *et al.* 1991

References:

Seiler A. *et al.*, Mol. Microbiol., 1996, 19(4):841-856.
Maiden R. *et al.*, Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.
Virji M. *et al.*, Mol. Microbiol., 1992, 6:1271-1279
Dempsey J.F. *et al.*, J. Bacteriol., 1991, 173:5476-5486

The amino acid sequences for each listed strain are as follows:

287_14 <SEQ ID 3179>
MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSKPAAPVVSEKETEAKEDAPQAGSQG
QGAPSAQGGQDMAAVSEENTGNGGAAATDKPKNEDEGAQNDMPQNAADTDSLTPNHTPAS
NMPAGNMENQAPDAGESEQPANQPDMANTADGMQGDDPSAGGENAGNTAAQGTNQAENNQ
TAGSQNPASSTNPSATNSGGDFGRTNVGNSVVIDGPSQNITLTHCKGDSCSGNNFLDEEV
QLKSEFEKLSDADKISNYKKDGKNDGKNDKFVGLVADSVQMKGINQYIIFYKPKPTSFAR
FRRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNYRYLTYGAEKLP
GGSYALRVQGEPSKGEMLAGTAVYNGEVLHFHTENGRPSPSRGRFAAKVDFGSKSVDGII
DSGDGLHMGTQKFKAAIDGNGFKGTWTENGGGDVSGKFYGPAGEEVAGKYSYRPTDAEKG
GFGVFAGKKEQD*

287_2 <SEQ ID 3180>
MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSKPAAPVVSEKETEAKEDAPQAGSQG
QGAPSAQGGQDMAAVSEENTGNGGAAATDKPKNEDEGAQNDMPQNAADTDSLTPNHTPAS
NMPAGNMENQAPDAGESEQPANQPDMANTADGMQGDDPSAGGENAGNTAAQGTNQAENNQ
TAGSQNPASSTNPSATNSGGDFGRTNVGNSVVIDGPSQNITLTHCKGDSCSGNNFLDEEV
QLKSEFEKLSDADKISNYKKDGKNDGKNDKFVGLVADSVQMKGINQYIIFYKPKPTSFAR
FRRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNYRYLTYGAEKLP
GGSYALRVQGEPSKGEMLAGTAVYNGEVLHFHTENGRPSPSRGRFAAKVDFGSKSVDGII
DSGDGLHMGTQKFKAAIDGNGFKGTWTENGGGDVSGKFYGPAGEEVAGKYSYRPTDAEKG
GFGVFAGKKEQD*

287_21. <SEQ ID 3181>
MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSKPAAPVVSEKETEAKEDAPQAGSQG
QGAPSAQGSQDMAAVSEENTGNGGAVTADNPKNEDEVAQNDMPQNAAGTDSSTPNHTPDP
NMLAGNMENQATDAGESSQPANQPDMANAADGMQGDDPSAGGQNAGNTAAQGANQAGNNQ
AAGSSDPIPASNPAPANGGSNFGRVDLANGVLIDGPSQNITLTHCKGDSCSGNNFLDEEV
QLKSEFEKLSDADKISNYKKDGKNDKFVGLVADSVQMKGINQYIIFYKPKPTSFARFRRS

ARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNYRYLTYGAEKLPGGSY
ALRVQGEPAKGEMLAGAAVYNGEVLHFHTENGRPYPTRGRFAAKVDFGSKSVDGIIDSGD
DLHMGTQKFKAAIDGNGFKGTWTENGSGDVSGKFYGPAGEEVAGKYSYRPTDAEKGGFGV
FAGKKEQD*

287_9 <SEQ ID 3182>
MFKRSVIAMACIVALSACGGGGGGSPDVKSADTLSKPAAPVVTEDVGEEVLPKEKKDEEA
VSGAPQADTQDATAGKGGQDMAAVSAENTGNGGAATTDNPENKDEGPQNDMPQNAADTDS
STPNHTPAPNMPTRDMGNQAPDAGESAQPANQPDMANAADGMQGDDPSAGENAGNTADQA
ANQAENNQVGGSQNPASSTNPNATNGGSDFGRINVANGIKLDSGSENVTLTHCKDKVCDR
DFLDEEAPPKSEFEKLSDEEKINKYKKDEQRENFVGLVADRVEKNGTNKYVIIYKDKSAS
SSSARFRRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNYRYLTYG
AEKLSGGSYALSVQGEPAKGEMLAGTAVYNGEVLHFHMENGRPSPSGGRFAAKVDFGSKS
VDGIIDSGDDLHMGTQKFKAVIDGNGFKGTWTENGGGDVSGRFYGPAGEEVAGKYSYRPT
DAEKGGFGVFAGKKEQD*

FA1090 <SEQ ID 3183>
MFKRSVIAMACIFPLSACGGGGGGSPDVKSADTPSKPAAPVVAENAGEGVLPKEKKDEEA
AGGAPQADTQDATAGEGSQDMAAVSAENTGNGGAATTDNPKNEDAGAQNDMPQNAAESAN
QTGNNQPAGSSDSAPASNPAPANGGSDFGRTNVGNSVVIDGPSQNITLTHCKGDSCNGDN
LLDEEAPSKSEFEKLSDEEKIKRYKKDEQRENFVGLVADRVKKDGTNKYIIFYTDKPPTR
SARSRRSLPAEIPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNYRYLTYGAEKLPGGS
YALRVQGEPAKGEMLVGTAVYNGEVLHFHMENGRPYPSGGRFA[illegible]FGSKSVDGIIDSG
DDLHMGTQKFKAAIDGNGFKGTWTENGGGDVSGRFYGPAGEEVAGKYSYRPTDAEKGGFG
VFAGKKDRD*

Z2491 <SEQ ID 3184>
MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSKPAAPVVSEKETEAKEDAPQAGSQG
QGAPSAQGSQDMAAVSEENTGNGGAVTADNPKNEDEVAQNDMPQNAAGTDSSTPNHTPDP
NMLAGNMENQATDAGESSQPANQPDMANAADGMQGDDPSAGGQNAGNTAAQGANQAGNNQ
AAGSSDPIPASNPAPANGGSNFGRVDLANGVLIDGPSQNITLTHCKGDSCSGNNFLDEEV
QLKSEFEKLSDADKISNYKKDGKNDKFVGLVADSVQMKGINQYIIFYKPKPTSFARFRRS
ARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNYRYLTYGAEKLPGGSY
ALRVQGEPAKGEMLAGAAVYNGEVLHFHTENGRPYPTRGRFAAKVDFGSKSVDGIIDSGD
DLHMGTQKFKAAIDGNGFKGTWTENGSGDVSGKFYGPAGEEVAGKYSYRPTDAEKGGFGV
FAGKKEQD*

Figure 21 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 287, further confirming its utility as an antigen for both vaccines and diagnostics.

EXAMPLE 14

Table 5 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 519 among different strains.

Table 5

519 gene variability: **List of used Neisseria strains**	
Identification Strains	**Source / reference**

number		
Group B		
zv01_519	NG6/88	R. Moxon / Seiler *et al.*, 1996
zv02_519	BZ198	R. Moxon / Seiler *et al.*, 1996
zv03_519ass	NG3/88	R. Moxon / Seiler *et al.*, 1996
zv04_519	297-0	R. Moxon / Seiler *et al.*, 1996
zv05_519	1000	R. Moxon / Seiler *et al.*, 1996
zv06_519ass	BZ147	R. Moxon / Seiler *et al.*, 1996
zv07_519	BZ169	R. Moxon / Seiler *et al.*, 1996
zv11_519	NGE31	R. Moxon / Seiler *et al.*, 1996
zv12_519	NGF26	R. Moxon / Seiler *et al.*, 1996
zv18_519	BZ232	R. Moxon / Seiler *et al.*, 1996
zv19_519	BZ83	R. Moxon / Seiler *et al.*, 1996
zv20_519ass	44/76	R. Moxon / Seiler *et al.*, 1996
zv21_519ass	MC58	R. Moxon
zv96_519	2996	Our collection
Group A		
zv22_519ass	205900	R. Moxon
z2491_519	Z2491	R. Moxon / Maiden *et al.*, 1998
Others		
zv26_519	A22	(group W) R. Moxon / Maiden *et al.*, 1998
zv27_519	E26	(group X) R. Moxon / Maiden *et al.*, 1998
zv28_519	860800	(group Y) R. Moxon / Maiden *et al.*, 1998
zv29_519ass	E32	(group Z) R. Moxon / Maiden *et al.*, 1998
Gonococcus		
zv32_519	Ng F62	R. Moxon / Maiden *et al.*, 1998
fa1090_519	FA1090	R. Moxon

References:

Seiler A. *et al.*, Mol. Microbiol., 1996, 19(4):841-856.
Maiden et al., Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.

The amino acid sequences for each listed strain are as follows:

```
FA1090_519 <SEQ ID 3185>
MEFFIILLAAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRAMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*
```

Z2491_519 <SEQ ID 3186>
MEFFIILLAAVVVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV01_519 <SEQ ID 3187>
MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV02_519 <SEQ ID 3188>
MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV03_519 <SEQ ID 3189>
MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV04_519 <SEQ ID 3190>
MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV05_519 <SEQ ID 3191>
MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV06_519ASS <SEQ ID 3192>
MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVFSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERK
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV07_519 <SEQ ID 3193>
MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV11_519 <SEQ ID 3194>
MEFFIILLAAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV12_519 <SEQ ID 3195>
MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV18_519 <SEQ ID 3196>
MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV19_519 <SEQ ID 3197>
MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV20_519ASS <SEQ ID 3198>
MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSM
ISAGMKIIDSSKTAK*

ZV21_519ASS <SEQ ID 3199>
MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV22_519ASS <SEQ ID 3200>
MEFFIILLAAVVVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAKIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV26_519 <SEQ ID 3201>
MEFFIILLAAVVVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV27_519 <SEQ ID 3202>

MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV28_519 <SEQ ID 3203>
MEFFIILLAAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV29_519ASS <SEQ ID 3204>
MEFFIILLAAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSIVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREPEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSNKTAK*

ZV32_519 <SEQ ID 3205>
MEFFIILLAAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRAMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV96_519 <SEQ ID 3206>
MEFFIILLAAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

Figure 22 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 519, further confirming its utility as an antigen for both vaccines and diagnostics.

EXAMPLE 15

Table 6 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 919 among different strains.

Table 6

919 gene variability: List of used Neisseria strains	
Identification Strains	**Source / reference**

number		
	Group B	
zm01	NG6/88	R. Moxon / Seiler *et al.*, 1996
zm02	BZ198	R. Moxon / Seiler *et al.*, 1996
zm03	NG3/88	R. Moxon / Seiler *et al.*, 1996
zm04	297-0	R. Moxon / Seiler *et al.*, 1996
zm05	1000	R. Moxon / Seiler *et al.*, 1996
zm06	BZ147	R. Moxon / Seiler *et al.*, 1996
zm07	BZ169	R. Moxon / Seiler *et al.*, 1996
zm08n	528	R. Moxon / Seiler *et al.*, 1996
zm09	NGP165	R. Moxon / Seiler *et al.*, 1996
zm10	BZ133	R. Moxon / Seiler *et al.*, 1996
zm11asbc	NGE31	R. Moxon / Seiler *et al.*, 1996
zm12	NGF26	R. Moxon / Seiler *et al.*, 1996
zm13	NGE28	R. Moxon / Seiler *et al.*, 1996
zm14	NGH38	R. Moxon / Seiler *et al.*, 1996
zm15	SWZ107	R. Moxon / Seiler *et al.*, 1996
zm16	NGH15	R. Moxon / Seiler *et al.*, 1996
zm17	NGH36	R. Moxon / Seiler *et al.*, 1996
zm18	BZ232	R. Moxon / Seiler *et al.*, 1996
zm19	BZ83	R. Moxon / Seiler *et al.*, 1996
zm20	44/76	R. Moxon / Seiler *et al.*, 1996
zm21	MC58	R. Moxon
zm96	2996	Our collection
	Group A	
zm22	205900	R. Moxon
zm23asbc	F6124	R. Moxon
z2491	Z2491	R. Moxon / Maiden *et al.*, 1998
	Group C	
zm24	90/18311	R. Moxon
zm25	93/4286	R. Moxon
	Others	
zm26	A22 (group W)	R. Moxon / Maiden *et al.*, 1998
zm27bc	E26 (group X)	R. Moxon / Maiden *et al.*, 1998
zm28	860800 (group Y)	R. Moxon / Maiden *et al.*, 1998
zm29asbc	E32 (group Z)	R. Moxon / Maiden *et al.*, 1998
zm31asbc	*N. lactamica*	R. Moxon
	Gonococcus	
zm32asbc	Ng F62	R. Moxon / Maiden *et al.*, 1998
zm33asbc	Ng SN4	R. Moxon
fa1090	FA1090	R. Moxon

References:

Seiler A. *et al.*, Mol. Microbiol., 1996, 19(4):841-856.
Maiden et al., Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.

The amino acid sequences for each listed strain are as follows:

FA1090 <SEQ ID 3207>
MKKHLLRSALYGIAAAILAACQSRSIQTFPQPDTSVINGPDRPAGIPDPAGTTVAGGGAV
YTVVPHLSMPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKRFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDGRRTERARFPIYGIPDDFISVPLPAGLRGGKN
LVRIRQTGKNSGTIDNAGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNL.IPYVSIGRYMADKGYL
KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSGNEGPVGALGTPLMGEYAGA
IDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

Z2491 <SEQ ID 3208>
MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSVQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKAYMQQNPQRLAEVLGQNPSYIFFRELTGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM01 <SEQ ID 3209>
MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM02 <SEQ ID 3210>
MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM03 <SEQ ID 3211>
MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM04 <SEQ ID 3212>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVAGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNAGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKAYMQQNPQRLAEVLGQNPSYIFFRELTGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM05 <SEQ ID 3213>
MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLSCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM06 <SEQ ID 3214>
MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL
KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM07 <SEQ ID 3215>
MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM08N <SEQ ID 3216>
MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM09 <SEQ ID 3217>
MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVAGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL
KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSGNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM10 <SEQ ID 3218>
MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVAGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL

KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSGNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM11ASBC <SEQ ID 3219>
MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSVQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNAGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL
KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM12 <SEQ ID 3220>
MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM13 <SEQ ID 3221>
MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAEQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM14 <SEQ ID 3222>
MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVAGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNAGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL
KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSRNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM15 <SEQ ID 3223>
MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDLAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNHQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL
KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSGNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM16 <SEQ ID 3224>
MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPGRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL
KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM17 <SEQ ID 3225>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL
KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM18 <SEQ ID 3226>
MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM19 <SEQ ID 3227>
MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM20 <SEQ ID 3228>
MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM21 <SEQ ID 3229>
MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM22 <SEQ ID 3230>
MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSVQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKAYMQQNPQRLAEVLGQNPSYIFFRELTGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM23ASBC <SEQ ID 3231>
MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNAGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL

KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTSKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGETAGK
MKEPGYVWQLLPNGMKPEYRP*

ZM24 <SEQ ID 3232>
MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVAGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL
KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSGNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM25 <SEQ ID 3233>
MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVAGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL
KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSGNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM26 <SEQ ID 3234>
MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSVQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKAYMQQNPQRLAEVLGQNPSYIFFRELTGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM27BC <SEQ ID 3235>
MKKYLFRAALYGISAAILAACQSKSIQTFPQPDTSVINGPDRPAGIPDPAGTTVAGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNAGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGETAGK
MKEPGYVWQLLPNGMKPEYRP*

ZM28 <SEQ ID 3236>
MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM29ASBC <SEQ ID 3237>
MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSGNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATTHPITRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM31ASBC <SEQ ID 3238>

MKKHLFRAALYGIAAAILAACQSKSIQTFPQPDTSIIKGPDRPAGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNAGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYVFFRELAGSGNDGFVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM32ASBC <SEQ ID 3239>
MKKHLLRSALYGIAAAILAACQSRSIQTFPQPDTSVINGPDRPAGIPDPAGTTVAGGGAV
YTVVPHLSMPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKRFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDGRRTERARFPIYGIPDDFISVPLPAGLRGGKA
LVRIRQTGKNSGTIDNAGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSGGDGPVGALGTPLMGGYAGA
IDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM33ASBC <SEQ ID 3240>
MKKHLLRSALYGIAAAILAACQSRSIQTFPQPDTSVINGPDRPAGIPDPAGTTVAGGGAV
YTVVPHLSMPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPIHSFQAKRFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDGRRTERARFPIYGIPDDFISVPLPAGLRGGKN
LVRIRQTGKNSGTIDNAGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMRQNPHKLAEVLGQNPSYIFFRELAGSGNEGPVGALGTPLMGEYAGA
IDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM96 <SEQ ID 3241>
MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

Figure 23 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 919, further confirming its utility as an antigen for both vaccines and diagnostics.

EXAMPLE 16

Using the above-described procedures, the following oligonucleotide primers were employed in the polymerase chain reaction (PCR) assay in order to clone the ORFs as indicated:

Table 7: Oligonucleotides used for PCR to amplify complete or partial ORFs

ORF	primer	Sequence	Restriction sites
001	Forward	CGCGGATCCCATATG-TGGATGGTGCTGGTCAT	BamHI-NdeI
	Reverse	CCCGCTCGAG-TGCCGTCTTGTCCCAC	XhoI
003	Forward	CGCGGATCCCATATG-GTCGTATTCGTGGC	BamHI-NdeI
	Reverse	CCCGCTCGAG-AAAATCATGAACACGCGC	XhoI
005	Forward	CGCGGATCCCATATG-GACAATATTGACATGT	BamHI-NdeI
	Reverse	CCCGCTCGAG-CATCACATCCGCCCG	XhoI
006	Forward	CGCGGATCCCATATG-CTGCTGGTGCTGG	BamHI-NdeI
	Reverse	CCCGCTCGAG-AGTTCCGGCTTTGATGT	XhoI
007	Forward	CGCGGATCCCATATG-GCCGACAACAGCATCAT	BamHI-NdeI
	Reverse	CCCGCTCGAG-AAGGCGTTCATGATATAAG	XhoI
008	Forward	CGCGGATCCCATATG-AACAACAGACATTTTG	BamHI-NdeI
	Reverse	CCCGCTCGAG-CCTGTCCGGTAAAAGAC	XhoI
009	Forward	CGCGGATCCCATATG-CCCCGCGCTGCT	BamHI-NdeI
	Reverse	CCCGCTCGAG-TGGCTTTTGCCACGTTTT	XhoI
011	Forward	CGCGGATCCCATATG-AAGACACACCGCAAG	BamHI-NdeI
	Reverse	CCCGCTCGAG-GGCGGTCAGTACGGT	XhoI
012	Forward	CGCGGATCCCATATG-CTCGCCCGTTGCC	BamHI-NdeI
	Reverse	CCCGCTCGAG-AGCGGGGAAGAGGCAC	XhoI
013	Forward	CGCGGATCCCATATG-CCTTTGACCATGCT	BamHI-NdeI
	Reverse	CCCGCTCGAG-CTGATTCGGCAAAAAAATCT	XhoI
018	Forward	CGCGGATCCCATATG-CAGCAGAGGCAGTT	BamHI-NdeI
	Reverse	CCCGCTCGAG-GACGAGGCGAACGCC	XhoI
019	Forward	AAAGAATTC-CTGCCAGCCGGCAAGACCCCGGC	Eco RI
	Reverse	AAACTGCAG-TCAGCGGGCGGGGACAATGCCCAT	Pst I
023	Forward	AAAGAATTC-AAAGAATATTCGGCATGGCAGGC	Eco RI
	Reverse	AAACTGCAG-TTACCCCCAAATCACTTTAACTGA	Pst I
025	Forward	AAAGAATTC-TGCGCCACCCAACAGCCTGCTCC	Eco RI
	Reverse	AAACTGCAG-TCAGAACGCGATATAGCTGTTCGG	Pst I
031	Forward	CGCGGATCCCATATG-GTCTCCCTTCGCTT	BamHI-NdeI
	Reverse	CCCGCTCGAG-ATGTAAGACGGGGACAAC	XhoI
032	Forward	CGCGGATCCCATATG-CGGCGAAACGTGC	BamHI-

			NdeI
	Reverse	CCCGCTCGAG-CTGGTTTTTGATATTTGTG	XhoI
033	Forward	CGCGGATCCCATATG-GCGGCGGCAGACA	BamHI-NdeI
	Reverse	CCCGCTCGAG-ATTTGCCGCATCCCGAT	XhoI
034	Forward	CGCGGATCCCATATG-GCCGAAAACAGCTACGG	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTGACGATTTGGTTCAATT	XhoI
036	Forward	CGCGGATCCCATATG-CTGAAGCCGTGCG	BamHI-NdeI
	Reverse	CCCGCTCGAG-CCGGACTGCGTATCGG	XhoI
038	Forward	CGCGGATCCCATATG-ACCGATTTCCGCCA	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTCTACGCCGTACTGCC	XhoI
039	Forward	CGCGGATCCCATATG-CCGTCCGAACCGC	BamHI-NdeI
	Reverse	CCCGCTCGAG-TAGGATGACGAGGTAGG	XhoI
041	Forward	CGCGGATCCCATATG-TTCGTGCGCGAACCGC	BamHI-NdeI
	Reverse	CCCGCTCGAG-GCCCAAAAACTCTTTCAAA	XhoI
042	Forward	CGCGGATCCCATATG-ACGATGATTTGCTTGC	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTGCAGCCTGCATTTGAC	XhoI
043	Forward	AAAAAAGGTACC-ATGGTTGTTTCAAATCAAAATATC	Kpn I
	Reverse	AAACTGCAG-TTATTGCGCTTCACCTTCCGCCGC	Pst I
043a	Forward	AAAAAAGGTACC-GCAAAAGTGCATGGCGGCTTGGACGGTGC	Kpn I
	Reverse	AAAAAACTGCAG-TTAATCCTGCAACACGAATTCGCCCGTCCG	Pst I
044	Forward	CGCGGATCCCATATG-CCGTCCGACTAGAG	BamHI-NdeI
	Reverse	CCCGCTCGAG-ATGCGCTACGGTAGCCA	XhoI
046	Forward	AAAGAATTC-ATGTCGGCAATGCTCCCGACAAG	Eco RI
	Reverse	AAACTGCAG-TCACTCGGCGACCCACACCGTGAA	Pst I
047	Forward	CGCGGATCCCATATG-GTCATCATACAGGCG	BamHI-NdeI
	Reverse	CCCGCTCGAG-TCCGAAAAAGCCCATTTTG	XhoI
048	Forward	AAAGAATTC-ATGCTCAACAAAGGCGAAGAATTGCC	Eco RI
	Reverse	AAACTGCAG-TCAAGATTCGACGGGGATGATGCC	Pst I
049	Forward	AAAGAATTC-ATGCGGGCGCAGGCGTTTGATCAGCC	Eco RI
	Reverse	AAACTGCAG-AAGGCGTATCTGAAAAAATGGCAG	Pst I
050	Forward	CGCGGATCCCATATG-GGCGCGGGCTGG	BamHI-NdeI
	Reverse	CCCGCTCGAG-AATCGGGCCATCTTCGA	XhoI
052	Forward	AAAAAAGAATTC-ATGGCTTTGGTGGCGGAGGAAAC	Eco RI
	Reverse	AAAAAAGTCGAC-TCAGGCGGCGTTTTTCACCTTCCT	Sal I
052a	Forward	AAAAAAGAATTC-GTGGCGGAGGAAACGGAAATATCCGC	Eco RI

	Reverse	AAAAAACTGCAG-TTAGCTGTTTTTGGAAACGCCGTCCAACCC	Pst I
073	Forward	CGCGGATCCCATATG-TGTATGCCATATAAGAT	BamHI-NdeI
	Reverse	CCCGCTCGAG-CACCGGATTGTCCGAC	XhoI
075	Forward	CGCGGATCCCATATG-CCGTCTTACTTCATC	BamHI-NdeI
	Reverse	CCCGCTCGAG-ATCACCAATGCCGATTATTT	XhoI
077a	Forward	AAAAAAGAATTC-GGCGGCATTTTCATCGACACCTTCCT	Eco RI
	Reverse	AAAAAACTGCAG-TCAGACGAACATCTGCACAAACGCAAT	Pst I
080	Forward	AAAGAATTC-GCGTCCGGGCTGGTTTGGTTTTACAATTC	Eco RI
	Reverse	AAACTGCAG-CTATTCTTCGGATTCTTTTTCGGG	Pst I
081	Forward	AAAGAATTC-ATGAAACCACTGGACCTAAATTTCATCTG	Eco RI
	Reverse	AAACTGCAG-TCACTTATCCTCCAATGCCTC	Pst I
082	Forward	AAAGAATTC-ATGTGGTTGTTGAAGTTGCCTGC	Eco RI
	Reverse	AAACTGCAG-TTACGCGGATTCGGCAGTTGG	Pst I
084	Forward	AAAGAATTC-TATCACCCAGAATATGAATACGGCTACCG	Eco RI
	Reverse	AAACTGCAG-TTATACTTGGGCGCAACATGA	Pst I
085	Forward	CGCGGATCCCATATG-GGTAAAGGGCAGGACT	BamHI-NdeI
	Reverse	CCCGCTCGAG-CAAAGCCTTAAACGCTTCG	XhoI
086	Forward	AAAAAAGGTACC-TATTTGGCATCAAAAGAAGGCGG	Kpn I
	Reverse	AAACTGCAG-TTACTCCACCCGATAACCGCG	Pst I
087	Forward	AAAGAATTC-ATGGGCGGTAAAACCTTTATGC	Eco RI
	Reverse	AAACTGCAG-TTACGCCGCACACGCAATCGC	Pst I
087a	Forward	AAAAAAGAATTC-AAGCTATTAGGCGTGCCGATTGTGATTCA	Eco RI
	Reverse	AAAAAACTGCAG-TTACGCCTGCAAGATGCCCAGCTTGCC	Pst I
088	Forward	AAAAAAGAATTC-ATGTTTTTATGGCTCGCACATTTCAG	Eco RI
	Reverse	AAAAAACTGCAG-TCAGCGGATTTTGAGGGTACTCAAACC	Pst I
089	Forward	CGCGGATCCCATATG-CCGCCCAAAATCAC	BamHI-NdeI
	Reverse	CCCGCTCGAG-TGCGCATACCAAAGCCA	XhoI
090	Forward	CGCGGATCCCATATG-CGCATAGTCGAGCA	BamHI-NdeI
	Reverse	CCCGCTCGAG-AGCAAAACGGCGGTACG	XhoI
091	Forward	AAAGAATTC-ATGGAAATACCCGTACCGCCGAGTCC	Eco RI
	Reverse	AAACTGCAG-TCAGCGCAGGGGGTAGCCCAAGCC	Pst I
092	Forward	AAAGAATTC-ATGTTTTTTATTTCAATCCG	Eco RI
	Reverse	AAACTGCAG-TCAAATCTGTTTCGACAATGC	Pst I
093	Forward	AAAGAATTC-ATGCAGAATTTTGGCAAAGTGGC	Eco RI
	Reverse	AAACTGCAG-CTATGGCTCGTCATACCGGGC	Pst I
094	Forward	AAAGAATTC-ATGCCGTCACGGAAGCGCATCAACTC	Eco RI
	Reverse	AAACTGCAG-TTATCCCGGCCATACCGCCGAACA	Pst I
095	Forward	AAAGAATTC-ATGTCCTTTCATTTGAACATGGACGG	Eco RI
	Reverse	AAACTGCAG-TCAACGCCGCAGGCACTAACGCCC	Pst I
096	Forward	AAAGAATTC-ATGGCTCGTCATACCGGGCAGGG	Eco RI

	Reverse	AAACTGCAG-TCAAAGGAAAAGGCCGTCTGAAAAGCG	Pst I
097	Forward	AAAGAATTC-ATGGACACTTCAAAACAAACACTGTTG	Eco RI
	Reverse	AAACTGCAG-TCAGCCCAAATACCAGAATTTCAG	Pst I
098	Forward	AAAGAATTC-GATGAACGCAGCCCAGCATGGATACG	Eco RI
	Reverse	AAACTGCAG-TTACGACATTCTGATTTGGCA	Pst I
102	Forward	AAAAAAGAATTC-GGCCTGATGATTTTGGAAGTCAACAC	Eco RI
	Reverse	AAAAAACTGCAG-TTATCCTTTAAATACGGGGACGAGTTC	Pst I
105	Forward	CGCGGATCCCATATG-TCCGCAAACGAATACG	BamHI-NdeI
	Reverse	CCCGCTCGAG-GTGTTCTGCCAGTTTCAG	XhoI
107	Forward	AAAAAAGAATTC-CTGATGATTTTGGAAGTCAACACCCATTATCC	Eco RI
	Reverse	AAAAAACTGCAG-TTATCCTTTAAATACGGGGACGAGTTC	Pst I
107b	Forward	AAAAAAGAATTC-GATACCCAAGCCCCCGCCGGCACAAACTACTG	Eco RI
	Reverse	AAAAAACTGCAG-TTACGCGTCGCCTTTAAAGTATTTGAGCAGGCTGGAGAC	Pst I
108	Forward	AAAGAATTC-ATGTTGCCGGGCTTCAACCG	Eco RI
	Reverse	AAACTGCAG-TTAGCGGTACAGGTGTTTGAAGCA	Pst I
108a	Forward	AAAAAAGAATTC-GGTAACACATTCGGCAGCTTAGACGGTGG	Eco RI
	Reverse	AAACTGCAG-TTAGCGGTACAGGTGTTTGAAGCA	Pst I
109	Forward	AAAGAATTC-ATGTATTATCGCCGGGTTATGGG	Eco RI
	Reverse	AAACTGCAG-CTAGCCCAAAGATTTGAAGTGTTC	Pst I
111	Forward	CGCGGATCCCATATG-TGTTCGGAACAAACCGC	BamHI-NdeI
	Reverse	CCCGCTCGAG-GCGGAGCAGTTTTTCAAA	XhoI
114	Forward	CGCGGATCCCATATG-GCTTCCATCACTTCGC	BamHI-NdeI
	Reverse	CCCGCTCGAG-CATCCGCGAAATCGTC	XhoI
117	Forward	AAAAAAGGTACC-ATGGTCGAAGAACTGGAACTGCTG	Kpn I
	Reverse	AAACTGCAG-TTAAAGCCGGGTAACGCTCAATAC	Pst I
118	Forward	AAAGTCGACATGTGTGAGTTCAAGGATATTATAAG	Sal I
	Reverse	AAAGCATGC-CTATTTTTTGTTGTAATAATCAAATC	Sph I
121	Forward	CGCGGATCCCATATG-GAAACACAGCTTTACAT	BamHI-NdeI
	Reverse	CCCGCTCGAG-ATAATAATATCCCGCGCCC	XhoI
122	Forward	CGCGGATCCCATATG-GTCATGATTAAAATCCGCA	BamHI-NdeI
	Reverse	CCCGCTCGAG-AATCTTGGTAGATTGGATTT	XhoI
125	Forward	AAAGAATTC-ATGTCGGGCAATGCCTCCTCTCC	Eco RI
	Reverse	AAACTGCAG-TCACGCCGTTTCAAGACG	Pst I
125a	Forward	AAAAAAGAATTC-ACGGCAGGCAGCACCGCCGCACAGGTTTC	Eco RI
	Reverse	AAAAAACTGCAG-TTATTTGCCACGTCGGTTTCTCCGGTGAACAACGC	Pst I
126	Forward	CGCGGATCCCATATG-CCGTCTGAAACCC	BamHI-

			NdeI
	Reverse	CCCGCTCGAG-ATATTCCGCCGAATGCC	XhoI
127	Forward	AAAGAATTC-ATGGAAATATGGAATATGTTGGACACTTG	Eco RI
	Reverse	AAACTGCAG-TTAAAGTGTTTCGGAGCCGGC	Pst I
127a	Forward	AAAAAAGAATTC-AAGGAACTGATTATGTGTCTGTCGGG	Eco RI
	Reverse	AAACTGCAG-TTAAAGTGTTTCGGAGCCGGC	Pst I
128	Forward	CGCGGATCCCATATG-ACTGACAACGCACT	BamHI-NdeI
	Reverse	CCCGCTCGAG-GACCGCGTTGTCGAAA	XhoI
130	Forward	CGCGGATCCCATATG-AAACAACTCCGCGA	BamHI-NdeI
	Reverse	CCCGCTCGAG-GAATTTTGCACCGGATTG	XhoI
132	Forward	AAAGAATTC-ATGGAACCCTTCAAAACCTTAATTTG	Eco RI
	Reverse	AAAAAACTGCAG-TCACCATGTCGGCATTTGAAAAAC	Pst I
134	Forward	CGCGGATCCCATATG-TCCCAAGAAATCCTC	BamHI-NdeI
	Reverse	CCCGCTCGAG-CAGTTTGACCGAATGTTC	XhoI
135	Forward	CGCGGATCCCATATG-AAATACAAAAGAATCGTATT	BamHI-NdeI
	Reverse	CCCGCTCGAG-AAATTCGGTCAGAAGCAGG	XhoI
137	Forward	AAAAAAGGTACC-ATGATTACCCATCCCCAATTCGATCC	Kpn I
	Reverse	AAAAAACTGCAG-TCAGTGCTGTTTTTCATGCCGAA	Pst I
137a	Forward	AAAAAAGAATTC-GGCCGCAAACACGGCATCGGCTTCCT	Eco RI
	Reverse	AAAAAACTGCAG-TTAAGCGGGATGACGCGGCAGCATACC	Pst I
138	Forward	AAAAAAGAATTC-AACTCAGGCGAAGGAGTGCTTGTGGC	Eco RI
	Reverse	AAAAAATCTAGA-TCAGTTTAGGGATAGCAGGCGTAC	Xba I
141	Forward	AAAGAATTC-ATGAGCTTCAAAACCGATGCCGAAATCGC	Eco RI
	Reverse	AAACTGCAG-TCAGAACAAGCCGTGAATCACGCC	Pst I
142	Forward	CGCGGATCCCATATG-CGTGCCGATTTCATG	BamHI-NdeI
	Reverse	CCCGCTCGAG-AAACTGCTGCACATGGG	XhoI
143	Forward	AAAAAAGAATTC-ATGCTCAGTTTCGGCTTTCTCGGCGTTCAGAC	Eco RI
	Reverse	AAAAAACTGCAG-TCAAACCCCGCCGTGTGTTTCTTTAAT	Pst I
144	Forward	AAAAAAGAATTC-GGTCTGATCGACGGGCGTGCCGTAAC	Eco RI
	Reverse	AAAAAATCTAGA-TCGGCATCGGCCGGCATATGTCCG	Xba I
146	Forward	AAAAAAGAATTC-CGCCAAGTCGTCATTGACCACGACAAAGTC	Eco RI
	Reverse	AAAAAACTGCAG-TTAGGCATCGGCAAATAGGAAACTGGG	Pst I
147	Forward	AAAAAAGAATTC-ACTGAGCAATCGGTGGATTTGGAAAC	Eco RI
	Reverse	AAAAAATCTAGA-TTAGGTAAAGCTGCGGCCCATTTGCGG	Xba I
148	Forward	AAAAAAGAATTC-ATGGCGTTAAAAACATCAAACTTGGAACACGC	Eco RI
	Reverse	AAAAAATCTAGA-TCAGCCCTTCATACAGCCTTCGTTTTG	Xba I
149	Forward	CGCGGATCCCATATG-CTGCTTGACAACAAAGT	BamHI-

			NdeI
	Reverse	CCCGCTCGAG-AAACTTCACGTTCACGCC	XhoI
150	Forward	CGCGGATCCCATATG-CAGAACACAAATCCG	BamHI-NdeI
	Reverse	CCCGCTCGAG-ATAAACATCACGCTGATAGC	XhoI
151	Forward	AAAAAAGAATTC-ATGAAACAAATCCGCAACATCGCCATCATCGC	Eco RI
	Reverse	AAAAAACTGCAG-TCAATCCAGCTTTTTAAAGTGGCGGCG	Pst I
152	Forward	AAAAAAGAATTC-ATGAAAAACAAAACCAAAGTCTGGGACCTCCC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGGACAGGAGCAGGATGGCGGC	Pst I
153	Forward	AAAAAAGAATTC-ATGGCGTTTGCTTACGGTATGAC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGTCATGTTTTTCCGTTTCATT	Pst I
153a	Forward	AAAAAAGAATTC-CGGACTTCGGTATCGGTTCCCCAGCATTG	Eco RI
	Reverse	AAAAAACTGCAG-TTACGCCGACGAAATACTCAGACTTTTCGG	Pst I
154	Forward	CGCGGATCCCATATG-ACTGACAACAGCCC	BamHI-NdeI
	Reverse	CCCGCTCGAG-TCGGCTTCCTTTCGGG	XhoI
155	Forward	AAAAAAGAATTC-ATGAAAATCGGTATCCCACGCGAGTC	Eco RI
	Reverse	AAAAAACTGCAG-TTACCCTTTCTTAAACATATTCAGCAT	Pst I
156	Forward	AAAAAAGAATTC-GCACAGCAAAACGGTTTTGAAGC	Eco RI
	Reverse	AAAAAACTGCAG-TCAAGCAGCCGCGACAAACAGCCC	Pst I
157	Forward	CGCGGATCCCATATG-AGGAACGAGGAAAAAC	BamHI-NdeI
	Reverse	CCCGCTCGAG-AAAACACAATATCCCCGC	XhoI
158	Forward	AAAAAAGAATTC-GCGGAGCAGTTGGCGATGGCAAATTCTGC	Eco RI
	Reverse	AAAAAATCTAGA-TTATCCACAGAGATTGTTTCCCAGTTC	Xba I
160	Forward	CGCGGATCCCATATG-GACATTCTGGACAAAC	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTTTGCCCGCCTTCTTT	XhoI
163	Forward	AAAAAAGGTACC-ACCGTGCCGGATCAGGTGCAGATGTG	Kpn I
	Reverse	AAAAAATCTAGA-TTACTCTGCCAATTCCACCTGCTCGTG	Xba I
163a	Forward	AAAAAAGAATTC-CGGCTGGTGCAGATAATGAGCCAGAC	Eco RI
	Reverse	AAAAAATCTAGA-TTACTCTGCCAATTCCACCTGCTCGTG	Xba I
164	Forward	CGCGGATCCCATATG-AACCGGACTTATGCC	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTGTTTCCGTCAAACTGC	XhoI
165	Forward	CGCGGATCCGCTAGC-GCTGAAGCGACAGACG	BamHI-NheI
	Reverse	CCCGCTCGAG-AATATCCAATACTTTCGCG	XhoI
206	Forward	CGCGGATCCCATATG-AAACACCGCCAACCGA	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTCTGTAAAAAAAGTATGTGC	XhoI
209	Forward	CGCGGATCCCATATG-CTGCGGCATTTAGGA	BamHI-NdeI

	Reverse	CCCGCTCGAG-TACCCCTGAAGGCAAC	XhoI
211	Forward	AAAAAAGAATTC-ATGTTGCGGGTTGCTGCTGC	Eco RI
	Reverse	AAAAAACTGCAG-CTATCCTGCGGATTGGCATTGAAA	Pst I
212	Forward	CGCGGATCCCATATG-GACAATCTCGTATGG	BamHI-NdeI
	Reverse	CCCGCTCGAG-AGGGGTTAGATCCTTCC	XhoI
215	Forward	CGCGGATCCCATATG-GCATGGTTGGGTCGT	BamHI-NdeI
	Reverse	CCCGCTCGAG-CATATCTTTTGTATCATAAATC	XhoI
216	Forward	CGCGGATCCCATATG-GCAATGGCAGAAAACG	BamHI-NdeI
	Reverse	CCCGCTCGAG-TACAATCCGTGCCGCC	XhoI
217	Forward	CGCGGATCCCATATG-GCGGATGACGGTGTG	BamHI-NdeI
	Reverse	CCCGCTCGAG-ACCCCGAATATCGAATCC	XhoI
218	Forward	CGCGGATCCCATATG-GTCGCGGTCGATC	BamHI-NdeI
	Reverse	CCCGCTCGAG-TAACTCATAGAATCCTGC	XhoI
219	Forward	CGCGGATCCGCTAGC-ACGGCAAGGTTAAG	BamHI-NheI
	Reverse	CCCGCTCGAG-TTTAAACCATCTCCTCAAAAC	XhoI
223	Forward	CGCGGATCCCATATG-GAATTCAGGCACCAAGTA	BamHI-NdeI
	Reverse	CCCGCTCGAG-GGCTTCCCGCGTGTC	XhoI
225	Forward	CGCGGATCCCATATG-GACGAGTTGACCAACC	BamHI-NdeI
	Reverse	CCCGCTCGAG-GTTCAGAAAGCGGGAC	XhoI
226	Forward	AAAGAATTC-CTTGCGATTATCGTGCGCACGCG	Eco RI
	Reverse	AAACTGCAG-TCAAAATCCCAAAACGGGGAT	Pst I
228	Forward	CGCGGATCCCATATG-TCGCAAGAAGCCAAACAG	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTGGCGGCATCTTTCAT	XhoI
229	Forward	CGCGGATCCCATATG-CAAGAGGTTTTGCCC	BamHI-NdeI
	Reverse	CCCGCTCGAG-ACACAATATAGCGGATGAAC	XhoI
230	Forward	CGCGGATCCCATATG-CATCCGGGTGCCGAC	BamHI-NdeI
	Reverse	CCCGCTCGAG-AAGTTTGGCGGCTTCGG	XhoI
232	Forward	AAAAAAGAATTC-ATGTACGCTAAAAAAGGCGGTTTGGG	Eco RI
	Reverse	AAAAAACTGCAG-TCAAGGTTTTTTCCTGATTGCCGCCGC	Pst I
232a	Forward	AAAAAAGAATTC-GCCAAGGCTGCCGATACACAAATTGA	Eco RI
	Reverse	AAAAAACTGCAG-TTAAACATTGTCGTTGCCGCCCAGATG	Pst I
233	Forward	CGCGGATCCCATATG-GCGGACAAACCCAAG	BamHI-NdeI
	Reverse	CCCGCTCGAG-GACGGCATTGAGCAG	XhoI
234	Forward	CGCGGATCCCATATG-GCCGTTTCACTGACCG	BamHI-

			NdeI
	Reverse	GCCCAAGCTT-ACGGTTGGATTGCCATG	Hind III
235	Forward	CGCGGATCCCATATG-GCCTGCCAAGTTCAAA	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTGGGCTGCTCTTC	XhoI
236	Forward	CGCGGATCCCATATG-GCGCGTTTCGCCTT	BamHI-NdeI
	Reverse	CCCGCTCGAG-ATGGGTCGCGCGCCGT	XhoI
238	Forward	CGCGGATCCGCTAGC-AACGGTTTGGATGCCCG	BamHI-NheI
	Reverse	CCCGCTCGAG-TTTGTCTAAGTTCCTGATATG	XhoI
239	Forward	CCGGAATTCTACATATG-CTCCACCATAAAGGTATTG	EcoRI-NdeI
	Reverse	CCCGCTCGAG-TGGTGAAGAGCGGTTTAG	XhoI
240	Forward	CGCGGATCCCATATG-GACGTTGGACGATTTC	BamHI-NdeI
	Reverse	CCCGCTCGAG-AAACGCCATTACCCGATG	XhoI
241	Forward	CCGGAATTCTACATATG-CCAACACGTCCAACT	EcoRI-NdeI
	Reverse	CCCGCTCGAG-GAATGCGCCTGTAATTAATC	XhoI
242	Forward	CGCGGATCCCATATG-ATCGGCAAACTTGTTG	BamHI-NdeI
	Reverse	GCCCAAGCTT-ACCGATACGGTCGCAG	HindIII
243	Forward	CGCGGATCCCATATG-ACGATTTTTCGATGCTGC	BamHI-NdeI
	Reverse	CCCGCTCGAG-CGACTTGGTTACCGCG	XhoI
244	Forward	CGCGGATCCCATATG-CCGTCTGAAGCCC	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTTTTCGGTAGGGGATTT	XhoI
246	Forward	CGCGGATCCCATATG-GACATCGGCAGTGC	BamHI-NdeI
	Reverse	CCCGCTCGAG-CCCGCGCTGCTGGAG	XhoI
247	Forward	CGCGGATCCCATATG-GTCGGATCGAGTTAC	BamHI-NdeI
	Reverse	CCCGCTCGAG-AAGTGTTCTGTTTGCGCA	XhoI
248	Forward	CGCGGATCCCATATG-CGCAAACAGAACACT	BamHI-NdeI
	Reverse	CCCGCTCGAG-CTCATCATTATTGCTAACA	XhoI
249	Forward	CGCGGATCCCATATG-AAGAATAATGATTGCTTC	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTCCCGACCTCCGAC	XhoI
251	Forward	CGCGGATCCCATATG-CGTGCTGCGGTAGT	BamHI-NdeI
	Reverse	CCCGCTCGAG-TACGAAAGCCGGTCGTG	XhoI
253	Forward	AAAAAAGAATTC-ATGATTGACAGGAACCGTATGCTGCG	Eco RI
	Reverse	AAAAAACTGCAG-TTATTGGTCTTTCAAACGCCCTTCCTG	Pst I

253a	Forward	AAAAAAGAATTC-AAAATCCTTTTGAAAACAAGCGAAAACGG	Eco RI
	Reverse	AAAAAACTGCAG-TTATTGGTCTTTCAAACGCCCTTCCTG	Pst I
254	Forward	AAAAAAGAATTC-ATGTATACAGGCGAACGCTTCAATAC	Eco RI
	Reverse	AAAAAATCTAGA-TCAGATTACGTAACCGTACACGCTGAC	Xba I
255	Forward	CGCGGATCCCATATG-GCCGCGTTGCGTTAC	BamHI-NdeI
	Reverse	CCCGCTCGAG-ATCCGCAATACCGACCAG	XhoI
256	Forward	CGCGGATCCGCTAGC-TTTTAACACCGCCGGAC	BamHI-NheI
	Reverse	CCCGCTCGAG-ACGCCTGTTTGTGCGG	XhoI
257	Forward	CGCGGATCCCATATG-GCGGTTTCTTTCCTG	BamHI-NdeI
	Reverse	CCCGCTCGAG-GCGCGTGAATATCGCG	XhoI
258	Forward	AAAAAAGAATTC-GATTATTTCTGGTGGATTGTTGCGTTCAG	Eco RI
	Reverse	AAAAAACTGCAG-CTACGCATAAGTTTTTACCGTTTTTGG	Pst I
258a	Forward	AAAAAAGAATTC-GCGAAGGCGGTGGCCCAAGGCGA	Eco RI
	Reverse	AAAAAACTGCAG-CTACGCATAAGTTTTTACCGTTTTTGG	Pst I
259	Forward	CGCGGATCCCATATG-GAAGAGCTGCCTCCG	BamHI-NdeI
	Reverse	CCCGCTCGAG-GGCTTTTCCGGCGTTT	XhoI
260	Forward	CGCGGATCCCATATG-GGTGCGGGTATGGT	BamHI-NdeI
	Reverse	CCCGCTCGAG-AACAGGGCGACACCCT	XhoI
261	Forward	AAAAAAGAATTC-CAAGATACAGCTCGGGCATTCGC	Eco RI
	Reverse	AAAAAACTGCAG-TCAAACCAACAAGCCTTGGTCACT	Pst I
263	Forward	CGCGGATCCCATATG-GCACGTTTAACCGTA	BamHI-NdeI
	Reverse	CCCGCTCGAG-GGCGTAAGCCTGCAATT	XhoI
264	Forward	AAAAAAGGTACC-GCCGACGCAGTGGTCAAGGCAGAA	Kpn I
	Reverse	AAACTGCAG-TCAGCCGGCGGTCAATACCGCCCG	Pst I
265	Forward	AAAAAAGAATTC-GCGGAGGTCAAGAGAAGGTGTTTG	Eco RI
	Reverse	AAAAAACTGCAG-TTACGAATACGTCGTCAAAATGGG	Pst I
266	Forward	AAAGAATTC-CTCATCTTTGCCAACGCCCCCTTC	Eco RI
	Reverse	AAACTGCAG-CTATTCCCTGTTGCGCGTGTGCCA	Pst I
267	Forward	AAAGAATTC-TTCTTCCGATTCGATGTTAATCG	Eco RI
	Reverse	AAACTGCAG-TTAGTAAAAACCTTTCTGCTTGGC	Pst I
269	Forward	AAAGAATTC-TGCAAACCTTGCGCCACGTGCCC	Eco RI
	Reverse	AAACTGCAG-TTACGAAGACCGCAACGAAAGGCAGAG	Pst I
269a	Forward	AAAAAAGAATTC-GACTTTATCCAAAACACGGCTTCGCC	Eco RI
	Reverse	AAACTGCAG-TTACGAAGACCGCAACGAAAGGCAGAG	Pst I
270	Forward	AAAGAATTC-GCCGTCAAGCTCGTTTTGTTGCAATG	Eco RI
	Reverse	AAACTGCAG-TTATTCGGCGGTAAATGCCGTCTG	Pst I
271	Forward	CGCGGATCCCATATG-CCTGTGTGCAGCTCGAC	BamHI-NdeI
	Reverse	CCCGCTCGAG-TCCCAGCCCCGTGGAG	XhoI

272	Forward	AAAGAATTC-ATGACCGCAAAGGAAGAACTGTTCGC	Eco RI
	Reverse	AAACTGCAG-TCAGAGCAGTTCCAAATCGGGGCT	Pst I
273	Forward	AAAGAATTC-ATGAGTCTTCAGGCGGTATTTATATACCC	Eco RI
	Reverse	AAACTGCAG-TTACGCGTAAGAAAAAACTGC	Pst I
274	Forward	CGCGGATCCCATATG-ACAGATTIGGTTACGGAC	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTGCTTTCAGTATTATTGAA	XhoI
276	Forward	AAAAAAGAATTC-ATGATTTTGCCGTCGTCCATCACGATGATGCG	Eco RI
	Reverse	AAAAAACTGCAG-CTACACCACCATCGGCGAATTTATGGC	Pst I
277	Forward	AAAAAAGAATTC-ATGCCCCGCTTTGAGGACAAGCTCGTAGG	Eco RI
	Reverse	AAAAAACTGCAG-TCATAAGCCATGCTTACCTTCCAACAA	Pst I
277a	Forward	AAAAAAGAATTC-GGGGCGGCGGCTGGGTTGGACGTAGG	Eco RI
	Reverse	AAAAAACTGCAG-TCATAAGCCATGCTTACCTTCCAACAA	Pst I
278	Forward	AAAAAAGGTACC-GTCAAAGTTGTATTAATCGGGCCTTTGCC	Kpn I
	Reverse	AAAAAACTGCAG-TCATTCAACCATATCAAATCTGCC	Pst I
278a	Forward	AAAAAAGAATTC-AAAACTCTCCTAATTCGTCATAGTCG	Eco RI
	Reverse	AAAAAACTGCAG-TCATTCAACCATATCAAATCTGCC	Pst I
279	Forward	CGCGGATCCCATATG-TTGCCTGCAATCACGATT	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTAGAAGCGGGCGGCAA	XhoI
280	Forward	AAAAAAGGTACC-GCCCCCCTGCCGGTTGTAACCAG	Kpn I
	Reverse	AAAAAACTGCAG-TTATTGCTTCATCGCGTTGGTCAAGGC	Pst I
281	Forward	AAAAAAGAATTC-GCACCCGTCGGCGTATTCCTCGTCATGCG	Eco RI
	Reverse	AAAAAATCTAGA-GGTCAGAATGCCGCCTTCTTTGCCGAG	Xba I
281a	Forward	AAAAAAGAATTC-TCCTACCACATCGAAATTCCTTCCGG	Eco RI
	Reverse	AAAAAATCTAGA-GGTCAGAATGCCGCCTTCTTTGCCGAG	Xba I
282	Forward	AAAAAAGAATTC-CTTTACCTTGACCTGACCAACGGGCACAG	Eco RI
	Reverse	AAAAAACTGCAG-TCAACCTGCCAGTTGCGGGAATATCGT	Pst I
283	Forward	CGCGGATCCCATATG-GCCGTCTTTACTTGGAAG	BamHI-NdeI
	Reverse	CCCGCTCGAG-ACGGCAGTATTTGTTTACG	XhoI
284	Forward	CGCGGATCCCATATG-TTTGCCTGCAAAAGAATCG	BamHI-NdeI
	Reverse	CCCGCTCGAG-CCGACTTTGCAAAAACTG	XhoI
286	Forward	CGCGGATCCCATATG-GCCGACCTTTCCGAAAA	BamHI-NdeI
	Reverse	CCCGCTCGAG-GAAGCGCGTTCCCAAG	XhoI
287	Forward	CCGGAATTCTAGCTAGC-CTTTCAGCCTGCGGG	EcoRI-NheI
	Reverse	CCCGCTCGAG-ATCCTGCTCTTTTTTGCC	XhoI
288	Forward	CGCGGATCCCATATG-CACACCGGACAGG	BamHI-NdeI
	Reverse	CCCGCTCGAG-CGTATCAAAGACTTGCGT	XhoI
290	Forward	CGCGGATCCCATATG-GCGGTTTGGGGCGGA	BamHI-

			NdeI
	Reverse	CCCGCTCGAG-TCGGCGCGGCGGGC	XhoI
292	Forward	CGCGGATCCCATATG-TGCGGGCAAACGCCC	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTGATTTTTGCGGATGATTT	XhoI
294	Forward	AAAAAAGAATTC-GTCTGGTCGATTCGGGTTGTCAGAAC	Eco RI
	Reverse	AAAAAACTGCAG-TTACCAGCTGATATAAAACATCGCTTT	Pst I
295	Forward	CGCGGATCCCATATG-AACCGGCCGGCCTCC	BamHI-NdeI
	Reverse	CCCGCTCGAG-CGATATTTGATTCCGTTGC	XhoI
297	Forward	AAAAAAGAATTC-GCATACATTGCTTCGACAGAGAG	Eco RI
	Reverse	AAAAAACTGCAG-TCAATCCGATTGCGACACGGT	Pst I
298	Forward	AAAAAAGAATTC-CTGATTGCCGTGTGGTTCAGCCAAAACCC	Eco RI
	Reverse	AAAAAACTGCAG-TCATGGCTGTGTACTTGATGGTTGCGT	Pst I
299	Forward	CGCGGATCCGCTAGC-CTACCTGTCGCCTCCG	BamHI-NheI
	Reverse	CCCGCTCGAG-TTGCCTGATTGCAGCGG	XhoI
302	Forward	AAAAAAGAATTC-ATGAGTCAAACCGATACGCAACG	Eco RI
	Reverse	AAAAAACTGCAG-TTAAGGTGCGGGATAGAATGTGGGCGC	Pst I
305	Forward	AAAAAAGGTACC-GAATTTTTACCGATTTCCAGCACCGGA	Kpn I
	Reverse	AAAAAACTGCAG-TCATTCCCAACTTATCCAGCCTGACAG	Pst I
305a	Forward	AAAAAAGGTACC-TCCCGTTCGGGCAGTACGATTATGGG	Kpn I
	Reverse	AAAAAACTGCAG-TTACAAACCGACATCATGCAGGGTGAA	Pst I
306	Forward	CGCGGATCCCATATG-TTTATGAACAAATTTTCCC	BamHI-NdeI
	Reverse	CCCGCTCGAG-CCGCATCGGCAGAC	XhoI
308	Forward	CGCGGATCCCATATG-TTAAATCGGGTATTTTATC	BamHI-NdeI
	Reverse	CCCGCTCGAG-ATCCGCCATTCCCTGC	XhoI
311	Forward	AAAAAAGGTACC-ATGTTCAGTTTTGGCTGGGTGTTT	Kpn I
	Reverse	AAACTGCAG-ATGTTCATATTCCCTGCCTTCGGC	Pst I
312	Forward	AAAAAAGGTACC-ATGAGTATCCCATCCGGCGAAATT	Kpn I
	Reverse	AAACTGCAG-TCAGTTTTTCATCGATTGAACCGG	Pst I
313	Forward	AAAAAAGAATTC-ATGGACGACCCGCGCACCTACGGATC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGCGGCTGCCGCCGATTTTGCT	Pst I
401	Forward	CGCGGATCCCATATG-AAGGCGGCAACACAGC	BamHI-NdeI
	Reverse	CCCGCTCGAG-CCTTACGTTTTTCAAAGCC	XhoI
402	Forward	AAAAAAGAATTC-GTGCCTCAGGCATTTTCATTTACCCTTGC	Eco RI
	Reverse	AAAAAATCTAGA-TTAAATCCCTCTGCCGTATTTGTATTC	Xba I
402a	Forward	AAAAAAGAATTC-AGGCTGATTGAAAACAAACACGG	Eco RI
	Reverse	AAAAAATCTAGA-TTAAATCCCTCTGCCGTATTTGTATTC	Xba I
406	Forward	CGCGGATCCCATATG-TGCGGGACACTGACAG	BamHI-NdeI
	Reverse	CCCGCTCGAG-AGGTTGTCCTTGTCTATG	XhoI

501	Forward	CGCGGATCCCATATG-GCAGGCGGAGATGGC	BamHI-NdeI
	Reverse	CCCGCTCGAG-GGTGTGATGTTCACCC	XhoI
502	Forward	CGCGGATCCCATATG-GTAGACGCGCTTAAGCA	BamHI-NdeI
	Reverse	CCCGCTCGAG-AGCTGCATGGCGGCG	XhoI
503	Forward	CGCGGATCCCATATG-TGTTCGGGGAAAGGCG	BamHI-NdeI
	Reverse	CCCGCTCGAG-CCGCGCATTCCTCGCA	XhoI
504	Forward	CGCGGATCCCATATG-AGCGATATTGAAGTGACG	BamHI-NdeI
	Reverse	GCCCAAGCTT-TGATTCAAGTCCTTGCCG	HindIII
505	Forward	CGCGGATCCCATATG-TTTCGTTTACAATTCAGG	BamHI-NdeI
	Reverse	CCCGCTCGAG-CGGCGTTTTATAGCGG	XhoI
510	Forward	CGCGGATCCCATATG-CCTTCGCGGACAC	BamHI-NdeI
	Reverse	CCCGCTCGAG-GCGCACTGGCAGCG	XhoI
512	Forward	CGCGGATCCCATATG-GGACATGAAGTAACGGT	BamHI-NdeI
	Reverse	CCCGCTCGAG-AGGAATAGCCTTTGACG	XhoI
515	Forward	CGCGGATCCCATATG-GAGGAAATAGCCTTCGA	BamHI-NdeI
	Reverse	CCCGCTCGAG-AAATGCCGCAAAGCATC	XhoI
516	Forward	CGCGGATCCCATATG-TGTACGTTGATGTTGTGG	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTGCGGGCGGCATC	XhoI
517	Forward	CGCGGATCCCATATG-GGTAAAGGTGTGGAAATA	BamHI-NdeI
	Reverse	CCCGCTCGAG-GTGCGCCCAGCCGT	XhoI
518	Forward	AAAGAATTC-GCTTTTTTACTGCTCCGACCGGAAGG	Eco RI
	Reverse	AAACTGCAG-TCAAATTTCAGACTCTGCCAC	Pst I
519	Forward	CGCGGATCCCATATG-TTCAAATCCTTTGTCGTCA	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTGGCGGTTTTGCTGC	XhoI
520	Forward	CGCGGATCCCATATG-CCTGCGCTTCTTTCA	BamHI-NdeI
	Reverse	CCCGCTCGAG-ATATTTACATTTCAGTCGGC	XhoI
521	Forward	CGCGGATCCCATATG-GCCAAAATCTATACCTGC	BamHI-NdeI
	Reverse	CCCGCTCGAG-CATACGCCCCAGTTCC	XhoI
522	Forward	CGCGGATCCCATATG-ACTGAGCCGAAACAC	BamHI-NdeI
	Reverse	GCCCAAGCTT-TTCTGATTTCAAATCGGCA	HindIII
523	Forward	CGCGGATCCCATATG-GCTCTGCTTTCCGCG	BamHI-NdeI

	Reverse	CCCGCTCGAG-AGGGTGTGTGATAATAAGAAG	XhoI
525	Forward	CGCGGATCCCATATG-GCCGAAATGGTTCAAATC	BamHI-NdeI
	Reverse	CCCGCTCGAG-GCCCGTGCATATCATAAA	XhoI
527	Forward	AAAGAATTC-TTCCCTCAATGTTGCCGTTTTCG	Eco RI
	Reverse	AAACTGCAG-TTATGCTAAACTCGAAACAAATTC	Pst I
529	Forward	CGCGGATCCGCTAGC-TGCTCCGGCAGCAAAAC	BamHI-NheI
	Reverse	GCCCAAGCTT-ACGCAGTTCGGAATGGAG	HindIII
530	Forward	CGCGGATCCCATATG-AGTGCGAGCGCGG	BamHI-NdeI
	Reverse	CCCGCTCGAG-ACGACCGACTGATTCCG	XhoI
531	Forward	AAAAAAGAATTC-TATGCCGCCGCCTACCAAATCTACGG	Eco RI
	Reverse	AAAAAACTGCAG-TTAAAACAGCGCCGTGCCGACGACAAG	Pst I
532	Forward	AAAAAAGAATTC-ATGAGCGGTCAGTTGGGCAAAGGTGC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGTGTTCCAAGTC[illegible]TCGGTATCAAA	Pst I
532a	Forward	AAAAAAGAATTC-TTGGGTGTCGCGTTTGAGCCGGAAGT	Eco RI
	Reverse	AAAAAACTGCAG-TCAGTGTTCCAAGTGGTCGGTATCAAA	Pst I
535	Forward	AAAGAATTC-ATGCCCTTTCCCGTTTTCAGAC	Eco RI
	Reverse	AAACTGCAG-TCAGACGACCCCGCCTTCCCC	Pst I
537	Forward	CGCGGATCCCATATG-CATACCCAAAACCAATCC	BamHI-NdeI
	Reverse	CCCGCTCGAG-ATCCTGCAAATAAAGGGTT	XhoI
538	Forward	CGCGGATCCCATATG-GTCGAGCTGGTCAAAGC	BamHI-NdeI
	Reverse	CCCGCTCGAG-TGGCATTTCGGTTTCGTC	XhoI
539	Forward	CGCGGATCCGCTAGC-GAGGATTTGCAGGAAA	BamHI-NheI
	Reverse	CCCGCTCGAG-TACCAATGTCGGCAAATC	XhoI
542	Forward	AAAGAATTC-ATGCCGTCTGAAACCGTGTC	Eco RI
	Reverse	AAACTGCAG-TTACCGCGAACCGGTCAGGAT	Pst I
543	Forward	AAAAAAGAATTC-GCCTTCGATGGCGACGTTGTAGGTAC	Eco RI
	Reverse	AAAAAATCTAGA-TTAATGAAGAAGAACATATTGGAATTTTGG	Xba I
543a	Forward	AAAAAAGAATTC-GGCAAAACTCGTCATGAATTTGC	Eco RI
	Reverse	AAAAAATCTAGA-TTAATGAAGAAGAACATATTGGAATTTTGG	Xba I
544	Forward	AAAGAATTC-GCGCCCGCCTTCTCCCTGCCCGACCTGCACGG	Eco RI
	Reverse	AAACTGCAG-CTATTGCGCCACGCGCGTATCGAT	Pst I
544a	Forward	AAAAAAGAATTC-GCAAATGACTATAAAAACAAAAACTTCCAAGTACTTGC	Eco RI
	Reverse	AAACTGCAG-CTATTGCGCCACGCGCGTATCGAT	Pst I
547	Forward	AAAGAATTC-ATGTTCGTAGATAACGGATTTAATAAAAC	Eco RI
	Reverse	AAACTGCAG-TTAACAACAAAAAACAAACCGCTT	Pst I
548	Forward	AAAGAATTC-GCCTGCAAACCTCAAGACAACAGTGCGGC	Eco RI

	Reverse	AAACTGCAG-TCAGAGCAGGGTCCTTACATCGGC	Pst I
550	Forward	AAAAAAGTCGAC-ATGATAACGGACAGGTTTCATCTCTTTCATTTTCC	Sal I
	Reverse	AAACTGCAG-TTACGCAAACGCTGCAAAATCCCC	Pst I
550a	Forward	AAAAAAGAATTC-GTAAATCACGCCTTTGGAGTCGCAAACGG	Eco RI
	Reverse	AAACTGCAG-TTACGCAAACGCTGCAAAATCCCC	Pst I
552	Forward	AAAAAAGAATTC-TTGGCGCGTTGGCTGGATAC	Eco RI
	Reverse	AAACTGCAG-TTATTTCTGATGCCTTTTCCCAAC	Pst I
554	Forward	CGCGGATCCCATATG-TCGCCCGCGCCCAAC	BamHI-NdeI
	Reverse	CCCGCTCGAG-CTGCCCTGTCAGACAC	XhoI
556	Forward	AAAGAATTC-GCGGGCGGTTTGTTTGGACATCCCG	Eco RI
	Reverse	AAACTGCAG-TTAACGGTGCGGACGTTTCTGACC	Pst I
557	Forward	CGCGGATCCCATATG-TGCGGTTTCCACCTGAA	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTCCGCCTTCAGAAAGG	XhoI
558	Forward	AAAGAATTC-GAGCTTTATATGTTTCAACAGGGGACGGC	Eco RI
	Reverse	AAACTGCAG-CTAAACAATGCCGTCTGAAAGTGGAGA	Pst I
558a	Forward	AAAAAAGAATTC-ATTAGATTCTATCGCCATAAACAGACGGG	Eco RI
	Reverse	AAAAAACTGCAG-CTAAACAATGCCGTCTGAAAGTGGAGA	Pst I
560	Forward	AAAAAAGAATTC-TCGCCTTTCCGGGACGGGGCGCACAAGATGGC	Eco RI
	Reverse	AAAAAACTGCAG-TCATGCGGTTTCAGACGGCATTTTGGC	Pst I
561	Forward	CCGGAATTCTACATATG-ATACTGCCAGCCCGT	EcoRI-NdeI
	Reverse	CCCGCTCGAG-TTTCAAGCTTTCTTCAGATG	XhoI
562	Forward	CGCGGATCCCATATG-GCAAGCCCGTCGAG	BamHI-NdeI
	Reverse	CCCGCTCGAG-AGACCAACTCCAACTCGT	XhoI
565	Forward	CGCGGATCCCATATG-AAGTCGAGCGCGAAATAC	BamHI-NdeI
	Reverse	CCCGCTCGAG-GGCATTGATCGGCGGC	XhoI
566	Forward	CGCGGATCCCATATG-GTCGGTGGCGAAGAGG	BamHI-NdeI
	Reverse	CCCGCTCGAG-CGCATGGGCGAAGTCA	XhoI
567	Forward	CCGGAATTCTACATATG-AGTGCGAACATCCTTG	EcoRI-NdeI
	Reverse	CCCGCTCGAG-TTTCCCCGACACCCTCG	XhoI
568	Forward	CGCGGATCCCATATG-CTCAGGGTCAGACC	BamHI-NdeI
	Reverse	CCCGCTCGAG-CGGCGCGGCGTTCAG	XhoI
569	Forward	AAAAAAGAATTC-CTGATTGCCTTGTGGGAATATGCCCG	Eco RI
	Reverse	AAAAAACTGCAG-TTATGCATAGACGCTGATAACGGCAAT	Pst I
570	Forward	CGCGGATCCCATATG-GACACCTTCCAAAAAATCG	BamHI-NdeI
	Reverse	CCCGCTCGAG-GCGGGCGTTCATTTCTTT	XhoI

571	Forward	AAAAAAGAATTC-ATGGGTATTGCCGGCGCCGTAAATGTTTTGAACCC	Eco RI
	Reverse	AAAAAACTGCAG-TTATGGCCGACGCGCGGCTACCTGACG	Pst I
572	Forward	CGCGGATCCCATATG-GCGCAAAAAGGCAAAACC	BamHI-NdeI
	Reverse	CCCGCTCGAG-GCGCAGTGTGCCGATA	XhoI
573	Forward	CGCGGATCCCATATG-CCCTGTTTGTGCCG	BamHI-NdeI
	Reverse	CCCGCTCGAG-GACGGTGTCATTTCGCC	XhoI
574	Forward	CGCGGATCCCATATG-TGGTTTGCCGCCCGC	BamHI-NdeI
	Reverse	CCCGCTCGAG-AACTTCGATTTTATTCGGG	XhoI
575	Forward	CGCGGATCCCATATG-GTTTCGGGCGAGG	BamHI-NdeI
	Reverse	CCCGCTCGAG-CATTCCGAATCTGAACAG	XhoI
576	Forward	CGCGGATCCCATATG-GCCGCCCCCGCATCT	BamHI-NdeI
	Reverse	CCCGCTCGAG-ATTTACTTTTTTGATGTCGAC	XhoI
577	Forward	CGCGGATCCCATATG-GAAAGGAACGGTGTATTT	BamHI-NdeI
	Reverse	CCCGCTCGAG-AGGCTGTTTGGTAGATTCG	XhoI
578	Forward	CGCGGATCCCATATG-AGAAGGTTCGTACAG	BamHI-NdeI
	Reverse	CCCGCTCGAG-GCCAACGCCTCCACG	XhoI
579	Forward	CGCGGATCCCATATG-AGATTGGGCGTTTCCAC	BamHI-NdeI
	Reverse	CCCGCTCGAG-AGAATTGATGATGTGTATGT	XhoI
580	Forward	CGCGGATCCCATATG-AGGCAGACTTCGCCGA	BamHI-NdeI
	Reverse	CCCGCTCGAG-CACTTCCCCCGAAGTG	XhoI
581	Forward	CGCGGATCCCATATG-CACTTCGCCCAGC	BamHI-NdeI
	Reverse	CCCGCTCGAG-CGCCGTTTGGCTTTGG	XhoI
582	Forward	AAAAAAGAATTC-TTTGGAGAGACCGCGCTGCAATGCGC	Eco RI
	Reverse	AAAAAATCTAGA-TCAGATGCCGTCCCAGTCGTTGAA	Xba I
583	Forward	AAAAAAGAATTC-ACTGCCGGCAATCGACTGCATAATCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAACGGAGGTCAATATGATGAAATTG	Pst I
584	Forward	AAAAAAGAATTC-GCGGCTGAAGCATTGAATTACAATATTGTC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGAACTGAACCGTCCCATTGACGCT	Pst I
585	Forward	AAAAAAGGTACC-TCTTTCTGGCTGGTGCAGAACACCCTTGC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGTTCGCACTTTTTTCTGTTTTGGA	Pst I
586	Forward	CGCGGATCCCATATG-GCAGCCCATCTCG	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTCAGCGAATCAAGTTTC	XhoI
587	Forward	CGCGGATCCCATATG-GACCTGCCCTTGACGA	BamHI-

			NdeI
	Reverse	CCCGCTCGAG-AAATGTATGCTGTACGCC	XhoI
588	Forward	AAAAAAGAATTC-GCCGTCCTGACTTCCTATCAAGAACCAGG	Eco RI
	Reverse	AAAAAACTGCAG-TTATTTGTTTTTGGGCAGTTTCACTTC	Pst I
589	Forward	AAAAAAGAATTC- ATGCAACAAAAAATCCGTTTCCAAATCGAAGG	Eco RI
	Reverse	AAAAAACTGCAG-CTAATCGATTTTTACCCGTTTCAGGCG	Pst I
590	Forward	AAAAAAGAATTC-ATGAAAAAACCTTTGATTTCAGTTGCGGC	Eco RI
	Reverse	AAAAAACTGCAG-TTACTGCTGCGGCTCTGAAACCAT	Pst I
591	Forward	AAAAAAGAATTC-CACTACATCGTTGCCAGATTGTGCGG	Eco RI
	Reverse	AAAAAACTGCAG-CTAACCGAGCAGCCGGGTAACGTCGTT	Pst I
592a	Forward	AAAAAAGAATTC-CGCGATTACACCGCCAAGCTGAAAATGGG	Eco RI
	Reverse	AAAAAACTGCAG-TTACCAAACGTCGGATTTGATACG	Pst I
593	Forward	CGCGGATCCGCTAGC-CTTGAACTGAACGGACTC	BamHI- NheI
	Reverse	CCCGCTCGAG-GCGGAAGCGGACGATT	XhoI
594a	Forward	AAAAAAGAATTC-GGTAAGTTCGCCGTTCAGGCCTTTCA	Eco RI
	Reverse	AAAAAACTGCAG-TTACGCCGCCGTTTCCTGACACTCGCG	Pst I
595	Forward	AAAAAAGAATTC-TGCCAGCCGCCGGAGGCGGAGAAAGC	Eco RI
	Reverse	AAAAAACTGCAG-TTATTTCAAGCCGAGTATGCCGCG	Pst I
596	Forward	CGCGGATCCCATATG-TCCCAACAATACGTC	BamHI- NdeI
	Reverse	CCCGCTCGAG-ACGCGTTACCGGTTTGT	XhoI
597	Forward	CGCGGATCCCATATG-CTGCTTCATGTCAGC	BamHI- NdeI
	Reverse	GCCCAAGCTT-ACGTATCCAGCTCGAAG	HindIII
601	Forward	CGCGGATCCCATATG-ATATGTTCCCAACCGGCAAT	BamHI- NdeI
	Reverse	CCCGCTCGAG-AAAACAATCCTCAGGCAC	XhoI
602	Forward	CGCGGATCCGCTAGC-TTGCTCCATCAATGC	BamHI- NheI
	Reverse	CCCGCTCGAG-ATGCAGCTGCTAAAAGCG	XhoI
603	Forward	AAAAAAGAATTC-CTGTCCTCGCGTAGGCGGGGACGGGG	Eco RI
	Reverse	AAAAAACTGCAG-CTACAAGATGCCGGCAAGTTCGGC	Pst I
604	Forward	CGCGGATCCGCTAGC-CCCGAAGCGCACTT	BamHI- NheI
	Reverse	CCCGCTCGAG-GACGGCATCTGCACGG	XhoI
606a	Forward	AAAAAAGAATTC-CGCGAATACCGCGCCGATGCGGGCGC	Eco RI
	Reverse	AAAAAACTGCAG-TTAAAGCGATTTGAGGCGGGCGATACG	Pst I
607	Forward	AAAAAAGAATTC-ATGCTGCTCGACCTCAACCGCTTTTC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGACGGCCTTATGCGATCTGAC	Pst I
608	Forward	AAAAAAGAATTC-ATGTCCGCCCTCCTCCCCATCATCAACCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAGTCTATCCAAATGTCGCGTTC	Pst I
609	Forward	CGCGGATCCCATATG-GTTGTGGATAGACTCG	BamHI- NdeI

	Reverse	CCCGCTCGAG-CTGGATTATGATGTCTGTC	XhoI
610	Forward	CGCGGATCCCATATG-ATTGGAGGGCTTATGCA	BamHI-NdeI
	Reverse	CCCGCTCGAG-ACGCTTCAACATCTTTGCC	XhoI
611	Forward	CGCGGATCCCATATG-CCGTCTCAAAACGGG	BamHI-NdeI
	Reverse	CCCGCTCGAG-AACGACTTTGAACGCGCAA	XhoI
613	Forward	CGCGGATCCCATATG-TCGCGTTCGAGCCG3	BamHI-NdeI
	Reverse	CCCGCTCGAG-AGCCTGTAAAATAAGCGGC	XhoI
614	Forward	CGCGGATCCCATATG-TCCGTCGTGAGCGGC	BamHI-NdeI
	Reverse	CCCGCTCGAG-CCATACTGCGGCGTTC	XhoI
616	Forward	AAAAAAGAATTC-ATGTCAAACACAATCAAAATGGTTGTCGG	Eco RI
	Reverse	AAAAAATCTAGA-TTAGTCCGGGCGGCAGGCAGCTCG	Xba I
619a	Forward	AAAAAAGAATTC-GGGCTTCTCGCCGCCTCGCTTGC	Eco RI
	Reverse	AAAAAACTGCAG-TCATTTTTTGTGTTTTAAAACGAGATA	Pst I
622	Forward	CGCGGATCCCATATG-GCCGCCCTGCCTAAAG	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTGTCCAAATGATAAATCTG	XhoI
624	Forward	CGCGGATCCCATATG-TCCCCGCGCTTTTACCG	BamHI-NdeI
	Reverse	CCCGCTCGAG-AGATTCGGGCCTGCGC	XhoI
625	Forward	CGCGGATCCCATATG-TTTGCAACCAGGAAAATG	BamHI-NdeI
	Reverse	CCCGCTCGAG-CGGCAAAATTACCGCCTT	XhoI
627a	Forward	AAAAAAGAATTC-AAAGCAGGCGAGGCAGGCGCGCTGGG	Eco RI
	Reverse	AAAAAACTGCAG-TTACGAATGAAACAGGGTACCCGTCATCAAGGC	Pst I
628	Forward	AAAAAAGGTACC-GCCTTACAAACATGGATTTTGCGTTC	Kpn I
	Reverse	AAAAAACTGCAG-CTACGCACCTGAAGCGCTGGCAAA	Pst I
629a	Forward	AAAAAAGAATTC-GCCACCTTTATCGCGTATGAAAACGA	Eco RI
	Reverse	AAAAAACTGCAG-TTACAACACCGCCGTCCGGTTCAAACC	Pst I
630a	Forward	AAAAAAGAATTC-GCGGCTTTGGGTATTTCTTTCGG	Eco RI
	Reverse	AAAAAACTGCAG-TTAGGAGACTTCGCCAATGGAGCCGGG	Pst I
635	Forward	AAAAAAGAATTC-ATGACCCAGCGACGGGTCGGCAAGCAAAACCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAATCCACTATAATCCTGTTGCT	Pst I
638	Forward	AAAAAAGAATTC-ATGATTGGCGAAAAGTTTATCGTAGTTGG	Eco RI
	Reverse	AAAAAACTGCAG-TCACGAACCGATTATGCTGATCGG	Pst I
639	Forward	CGCGGATCCCATATG-ATGCTTTATTTTGTTCG	BamHI-NdeI
	Reverse	CCCGCTCGAG-ATCGCGGCTGCCGAC	XhoI
642	Forward	CGCGGATCCCATATG-CGGTATCCGCCGCAAT	BamHI-NdeI
	Reverse	CCCGCTCGAG-AGGATTGCGGGGCATTA	XhoI

643	Forward	CGCGGATCCCATATG-GCTTCGCCGTCGGCAG	BamHI-NdeI
	Reverse	CCCGCTCGAG-AACCGAAAAACAGACCGC	XhoI
644	Forward	AAAAAAGAATTC-ATGCCGTCTGAAAGGTCGGCGGATTGTTGCCC	Eco RI
	Reverse	AAAAAATCTAGA-CTACCCGCAATATCGGCAGTCCAATAT	Pst I
645	Forward	AAAAAAGAATTC-GTGGAACAGAGCAACACGTTAAATCG	Eco RI
	Reverse	AAAAAACTGCAG-CTACGAGGAAACCGAAGACCAGGCCGC	Pst I
647	Forward	AAAAAAGAATTC-ATGCAAAGGCTCGCCGCAGACGG	Eco RI
	Reverse	AAAAAACTGCAG-TTAGATTATCAGGGATATCCGGTAGAA	Pst I
648	Forward	AAAAAAGAATTC-ATGAACAGGCGCGACGCGCGGATCGAACG	Eco RI
	Reverse	AAAAAACTGCAG-TCAAGCTGTGTGCTGATTGAATGCGAC	Pst I
649	Forward	AAAAAAGAATTC-GGTACGTCAGAACCCGCCCACCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAACGGCGGAAACTGCCGCCGTC	Pst I
650	Forward	AAAAAAGAATTC-ATGTCCAAACTCAAAACCATCGC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGACGGCATGGCGGTCTGTTTT	Pst I
652	Forward	AAAAAAGGTACC-GCTGCCGAAGACTCAGGCCTGCCGCTTTACCG	Kpn I
	Reverse	AAAAAACTGCAG-TTATTTGCCCAGTTGGTAGAATGCGGC	Pst I
653	Forward	AAAAAAGAATTC-GCGGCTTTGCCGGTAATTTTCATCGG	Eco RI
	Reverse	AAAAAACTGCAG-CTATGCCGGTCTGGTTGCCGGCGGCGA	Pst I
656a	Forward	AAAAAAGAATTC-CGGCCGACGTCGTTGCGTCCTAAGTC	Eco RI
	Reverse	AAAAAACTGCAG-CTACGATTTCGGCGATTTCCACATCGT	Pst I
657	Forward	AAAAAAGAATTC-GCAGAATTTGCCGACCGCCATTTGTGCGC	Eco RI
	Reverse	AAAAAACTGCAG-TTATAGGGACTGATGCAGTTTTTTTGC	Pst I
658	Forward	CGCGGATCCCATATG-GTGTCCGGAATTGTG	BamHI-NdeI
	Reverse	CCCGCTCGAG-GGCAGAATGTTTACCGTT	XhoI
661	Forward	AAAAAAGAATTC-ATGCACATCGGCGGCTATTTTATCGACAACCC	Eco RI
	Reverse	AAAAAACTGCAG-TCACGACGTGTCTGTTCGCCGTCGGGC	Pst I
663	Forward	CGCGGATCCCATATG-TGTATCGAGATGAAATT	BamHI-NdeI
	Reverse	CCCGCTCGAG-GTAAAAATCGGGGCTGC	XhoI
664	Forward	CGCGGATCCCATATG-GCGGCTGGCGCGGT	BamHI-NdeI
	Reverse	CCCGCTCGAG-AAATCGAGTTTTACACCAC	XhoI
665	Forward	AAAAAAGAATTC-ATGAAATGGGACGAAACGCGCTTCGG	Eco RI
	Reverse	AAAAAACTGCAG-TCAATCCAAAATTTTGCCGACGATTTC	Pst I
666	Forward	AAAAAAGAATTC-AACTCAGGCGAAGGAGTGCTTGTGGC	Eco RI
	Reverse	AAAAAATCTAGA-TCAGTTTAGGGATAGCAGGCGTAC	Xba I
667	Forward	AAAAAAGAATTC-CCGCATCCGTTTGATTTCCATTTCGTATTCGTCCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAATGACACAATAGGCGCAAGTC	Pst I

669	Forward	AAAAAAGAATTC-ATGCGCCGCATCATTAAAAAACACCAGCC	Eco RI
	Reverse	AAAAAACTGCAG-TTACAGTATCCGTTTGATGTCGGC	Pst I
670a	Forward	AAAAAAGAATTC-AAAAACGCTTCGGGCGTTTCGTCTTC	Eco RI
	Reverse	AAAAAACTGCAG-TTAGGAGCTTTTGGAACGCGTCGGACTGGC	Pst I
671	Forward	CGCGGATCCCATATG-ACCAGCAGGGTAAC	BamHI-NdeI
	Reverse	CCCGCTCGAG-AGCAACTATAAAAACGCAAG	XhoI
672	Forward	CGCGGATCCCATATG-AGGAAAATCCGCACC	BamHI-NdeI
	Reverse	CCCGCTCGAG-ACGGGATAGGCGGTTG	XhoI
673	Forward	AAAAAAGAATTC-ATGGATATTGAAACCTTCCTTGCAGG	Eco RI
	Reverse	AAAAAACTGCAG-CTACAAACCCAGCTCGCGCAGGAA	Pst I
674	Forward	AAAAAAGAATTC-ATGAAAACAGCCCGCCGCCGTTCCCG	Eco RI
	Reverse	AAAAAACTGCAG-TCAACGGCGTTTGGGCTCGTCGGG	Pst I
675	Forward	CGCGGATCCCATATG-AACACCATCGCCCC	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTCTTCGTCTTCAAACTGT	XhoI
677a	Forward	AAAAAAGAATTC-AGACGGCATTCCCGATCAGTCGATTTTGA	Eco RI
	Reverse	AAAAAACTGCAG-TTACGTATGCGCGAAATCGACCGCCGC	Pst I
680	Forward	CGCGGATCCGCTAGC-ACGAAGGGCAGTTCGG	BamHI-NheI
	Reverse	CCCGCTCGAG-CATCAAAAACCTGCCGC	XhoI
681	Forward	AAAAAAGAATTC-ATGACGACGCCGATGGCAATCAGTGC	Eco RI
	Reverse	AAAAAACTGCAG-TTACCGTCTTCCGCAAAAAACAGC	Pst I
683	Forward	CGCGGATCCCATATG-TGCAGCACACCGGACAA	BamHI-NdeI
	Reverse	CCCGCTCGAG-GAGTTTTTTCCGCATACG	XhoI
684	Forward	CGCGGATCCCATATG-TGCGGTACTGTGCAAAG	BamHI-NdeI
	Reverse	CCCGCTCGAG-CTCGACCATCTGTTGCG	XhoI
685	Forward	CGCGGATCCCATATG-TGTTTGCTTAATAATAAACATT	BamHI-NdeI
	Reverse	CCCGCTCGAG-CTTTTTCCCCGCCGCA	XhoI
686	Forward	CGCGGATCCCATATG-TGCGGCGGTTCGGAAG	BamHI-NdeI
	Reverse	CCCGCTCGAG-CATTCCGATTCTGATGAAG	XhoI
687	Forward	CGCGGATCCCATATG-TGCGACAGCAAAGTCCA	BamHI-NdeI
	Reverse	CCCGCTCGAG-CTGCGCGGCTTTTTGTT	XhoI
690	Forward	CGCGGATCCCATATG-TGTTCTCCGAGCAAAGAC	BamHI-NdeI
	Reverse	CCCGCTCGAG-TATTCGCCCCGTGTTTGG	XhoI
691	Forward	CGCGGATCCCATATG-GCCACGGCTTATATCCC	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTGAGGCAGGAAGAAAG	XhoI

694	Forward	CGCGGATCCCATATG-TTGGTTTCCGCATCCGG	BamHI-NdeI
	Reverse	CCCGCTCGAG-TCTGCGTCGGTGCGGT	XhoI
695	Forward	CGCGGATCCCATATG-TTGCCTCAAACTCGTCCG	BamHI-NdeI
	Reverse	CCCGCTCGAG-TCGTTTGCGCACGGCT	XhoI
696	Forward	CGCGGATCCCATATG-TTGGGTTGCCGGCAGG	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTGATTGCCGCAATGATG	XhoI
700a	Forward	AAAAAAGAATTC-GCATCGACAGACGGTGTGTCGTGGAC	Eco RI
	Reverse	AAAAAACTGCAG-TTACGCTACCGGCACGACTTCCAAACC	Pst I
701	Forward	CGCGGATCCCATATG-AAGACTTGTTTGGATACTTC	BamHI-NdeI
	Reverse	CCCGCTCGAG-TGCCGACAACAGCCTC	XhoI
702	Forward	AAAAAAGAATTC-ATGCCGTGTTCCAAAGCCAGTTGGATTTC	Eco RI
	Reverse	AAAAAACTGCAG-TTAACCCCATTCCACCCGGAGAACCGA	Pst I
703	Forward	CGCGGATCCGCTAGC-CAAACGCTGGCAACCG	BamHI-NheI
	Reverse	CCCGCTCGAG-TTTTGCAGGTTTGATGTTTG	XhoI
704a	Forward	AAAAAAGAATTC-GCTTCTACCGGTACGCTGGCGCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAGTTTTGCCGGATAATATGGCGGGTGCG	Pst I
707	Forward	CGCGGATCCGCTAGC-GAAATTATTAACGATGCAGA	BamHI-NheI
	Reverse	CCCGCTCGAG-GAAACTGTAATTCAAGTTGA	XhoI
708	Forward	CGCGGATCCGCTAGC-CCTTTTAAGCCATCCAAAA	BamHI-NheI
	Reverse	CCCGCTCGAG-TTGACCGGTGAGGACG	XhoI
710	Forward	CGCGGATCCCATATG-GAAACCCACGAAAAAATC	BamHI-NdeI
	Reverse	CCCGCTCGAG-AACGGTTTCGGTCAG	XhoI
714	Forward	CGCGGATCCCATATG-AGCTATCAAGACATCTT	BamHI-NdeI
	Reverse	CCCGCTCGAG-GCGGTAGGTAAATCGGAT	XhoI
716	Forward	CGCGGATCCCATATG-GCCAACAAACCGGCAAG	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTAGAACCGCATTTGCC	XhoI
718	Forward	CGCGGATCCCATATG-GAGCCGATAATGGCAAA	BamHI-NdeI
	Reverse	CCCGCTCGAG-GGCGCGGGCATGGTCTTGTCC	XhoI
720	Forward	CGCGGATCCCATATG-AGCGGATGGCATACC	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTTGCATAGCTGTTGACCA	XhoI
723	Forward	CGCGGATCCCATATG-CGACCCAAGCCCC	BamHI-NdeI
	Reverse	CCCGCTCGAG-AATGCGAATCCGCCGCC	XhoI

725	Forward	CGCGGATCCCATATG-GTGCGCACGGTTAAA	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTGCTTATCCTTAAGGGTTA	XhoI
726	Forward	CGCGGATCCCATATG-ACCATCTATTTCAAAAAC	BamHI-NdeI
	Reverse	CCCGCTCGAG-GCCGATGTTTAGCGTCC	XhoI
728	Forward	CGCGGATCCCATATG-TTTTGGCTGGGAACGGG	BamHI-NdeI
	Reverse	CCCGCTCGAG-GTGAGAAAGGTCGCGC	XhoI
729	Forward	CGCGGATCCCATATG-TGCACCATGATTCCCCA	BamHI-NdeI
	Reverse	GCCCAAGCTT-TTTGTCGGTTTGGGTATC	HindIII
731	Forward	CGCGGATCCGCTAGC-GCCGTGCCGGAGG	BamHI-NheI
	Reverse	CCCGCTCGAG-ACGGGCGCGGCAG	XhoI
732	Forward	CCGGAATTCTACATATG-TCGAAACCTGTTTTTAAGAA	EcoRI-NdeI
	Reverse	CCCGCTCGAG-CTTCTTATCTTTTTTATCTTTC	XhoI
733	Forward	CGCGGATCCCATATG-GCCTGCGGCGGCAA	BamHI-NdeI
	Reverse	CCCGCTCGAG-TCGCTTGCCTCCTTTAC	XhoI
734	Forward	CGCGGATCCCATATG-GCCGATACTTACGGCTAT	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTGAGATTTTGAATCAAAGAG	XhoI
735	Forward	CGCGGATCCCATATG-AAGCAGCAGGCGGTCA	BamHI-NdeI
	Reverse	CCCGCTCGAG-ATTTCCGTAGCCGAGGG	XhoI
737	Forward	CGCGGATCCCATATG-CACCACGACGGACACG	BamHI-NdeI
	Reverse	CCCGCTCGAG-GTCGTCGCGGCGGGA	XhoI
739	Forward	CGCGGATCCCATATG-GCAAAAAAACCGAACA	BamHI-NdeI
	Reverse	CCCGCTCGAG-GAAGAGTTTGTCGAGAATT	XhoI
740	Forward	CGCGGATCCCATATG-GCCAATCCGCCCGAAG	BamHI-NdeI
	Reverse	CCCGCTCGAG-AAACGCGCCAAAATAGTG	XhoI
741	Forward	CGCGGATCCCATATG-TGCAGCAGCGGAGGG	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTGCTTGGCGGCAAGGC	XhoI
743	Forward	CGCGGATCCCATATG-GACGGTGTTGTGCCTGTT	BamHI-NdeI
	Reverse	CCCGCTCGAG-CTTACGGATCAAATTGACG	XhoI
745	Forward	CGCGGATCCCATATG-TTTTGGCAACTGACCG	BamHI-NdeI
	Reverse	CCCGCTCGAG-CAAATCAGATGCCTTTAGG	XhoI
746	Forward	CGCGGATCCCATATG-TCCGAAAACAAACAAAAC	BamHI-

				NdeI
	Reverse	CCCGCTCGAG-TTCATTCGTTACCTGACC		XhoI
747	Forward	CCGGAATTCTAGCTAGC-CTGACCCCTTGGG		EcoRI-NheI
	Reverse	GCCCAAGCTT-TTTTGATTTTAATTGACTATAGAAC		HindIII
749	Forward	CGCGGATCCCATATG-TGCCAGCCGCCG		BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTCAAGCCGAGTATGC		XhoI
750	Forward	CGCGGATCCCATATG-TGTTCGCCCGAACCTG		BamHI-NdeI
	Reverse	CCCGCTCGAG-CTTTTTCCCCGCCGCAA		XhoI
758	Forward	CGCGGATCCCATATG-AACAATCTGACCGTGTT		BamHI-NdeI
	Reverse	CCCGCTCGAG-TGGCTCAATCCTTTCTGC		XhoI
759	Forward	CGCGGATCCGCTAGC-CGCTTCACACACACCAC		BamHI-NheI
	Reverse	CCCGCTCGAG-CCAGTTGTAGCCTATTTTG		XhoI
763	Forward	CGCGGATCCCATATG-CTGCCTGAAGCATGGCG		BamHI-NdeI
	Reverse	CCCGCTCGAG-TTCCGCAAATACCGTTTCC		XhoI
764	Forward	CGCGGATCCCATATG-TTTTTCTCCGCCCTGA		BamHI-NdeI
	Reverse	CCCGCTCGAG-TCGCTCCCTAAAGCTTTC		XhoI
765	Forward	CGCGGATCCCATATG-TTAAGATGCCGTCCG		BamHI-NdeI
	Reverse	CCCGCTCGAG-ACGCCGACGTTTTTATTAA		XhoI
767	Forward	CGCGGATCCCATATG-CTGACGGAAGGGGAAG		BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTCTGTACAGCAGGGG		XhoI
768	Forward	CGCGGATCCCATATG-GCCCCGCAAAAACCCG		BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTCATCCCTTTTTTGAGC		XhoI
770	Forward	CGCGGATCCCATATG-TGCGGCAGCGGCGAA		BamHI-NdeI
	Reverse	CCCGCTCGAG-GCGTTTGTCGAGATTTTC		XhoI
771	Forward	CGCGGATCCCATATG-TCCGTATATCGCACCTTC		BamHI-NdeI
	Reverse	CCCGCTCGAG-CGGTTCTTTAGGTTTGAG		XhoI
772	Forward	CGCGGATCCCATATG-TTTGCGGCGTTGGTGG		BamHI-NdeI
	Reverse	CCCGCTCGAG-CAATGCCGACATCAAACG		XhoI
774	Forward	CGCGGATCCCATATG-TCCGTTTCACCCGTTCC		BamHI-NdeI
	Reverse	CCCGCTCGAG-TCGTTTGCGCACGGCT		XhoI
790	Forward	CGCGGATCCCATATG-GCAAGAAGGTCAAAAAC		BamHI-NdeI

	Reverse	CCCGCTCGAG-GGCGTTGTTCGGATTTCG	XhoI
900	Forward	CGCGGATCCCATATG-CCGTCTGAAATGCCG	BamHI-NdeI
	Reverse	CCCGCTCGAG-ATATGGAAAAGTCTGTTGTC	XhoI
901	Forward	CGCGGATCCCATATG-CCCGATTTTTCGATG	BamHI-NdeI
	Reverse	CCCGCTCGAG-AAAATGGAACAATACCAGG	XhoI
902	Forward. 2	CCGGAATTCTACATATG-TTGCACTTTCAAAGGATAATC	EcoRI-NdeI
	Reverse	CCCGCTCGAG-AAAAATGTACAATGGCGTAC	XhoI
903	Forward	CCGGAATTCTAGCTAGC-CAGCGTCAGCAGCACAT	EcoRI-NheI
	Reverse	CCCGCTCGAG-GAAACTGTAATTCAAGTTGAA	XhoI
904	Forward	AAAAAAGGTACC-ATGATGCAGCACAATCGTTTC	Kpn I
	Reverse	AAACTGCAG-TTAATATCGATAGGTTATATG	Pst I
904a	Forward	AAAAAAGAATTC-CGGCTCGGCATTGTGCAGATGTTGCA	Eco RI
	Reverse	AAACTGCAG-TTAATATCGATAGGTTATATG	Pst I
905	Forward	CGCGGATCCCATATG-AACAAAATATACCGCATC	BamHI-NdeI
	Reverse	CCCGCTCGAG-CCACTGATAACCGACAGAT	XhoI
907	Forward	CGCGGATCCCATATG-GGCGCGCAACGTGAG	BamHI-NdeI
	Reverse	CCCGCTCGAG-ACGCCACTGCCAGCG	XhoI
908	Forward	AAAGAATTC-GCAGAGTTAGTAGGCGTTAATAAAAATAC	Eco RI
	Reverse	AAACTGCAG-TTAATATGGTTTTGTCGTTCG	Pst I
909	Forward	CGCGGATCCCATATG-TGCGCGTGGGAAACTTAT	BamHI-NdeI
	Reverse	CCCGCTCGAG-TCGGTTTTGAAACTTTGGTTTT	XhoI
910	Forward	AAAGAATTC-GCATTTGCCGGCGACTCTGCCGAGCG	Eco RI
	Reverse	AAACTGCAG-TCAGCGATCGAGCTGCTCTTT	Pst I
911	Forward	AAAGAATTC-GCTTTCCGCGTGGCCGGCGGTGC	Eco RI
	Reverse	AAAAAACTGCAG-GTCGACTTATTCGGCGGCTTTTTCCGC	Pst I
912	Forward	AAAAAAGAATTC-CAAATCCGTCAAAACGCCACTCAAGTATTGAG	Eco RI
	Reverse	AAAAAACTGCAG-TTACAGTCCGTCCACGCCTTTCGC	Pst I
913	Forward	CGCGGATCCCATATG-GAAACCCGCCCCGC	BamHI-NdeI
	Reverse	CCCGCTCGAG-AGGTTGTGTTCCAGGTTG	XhoI
915	Forward	CGCGGATCCCATATG-TGCCGGCAGGCGGAA	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTGAAAATATAGGTATCAGG	XhoI
914	Forward	AAAGAATTC-GACAGAATCGGCGATTTGGAAGCACG	Eco RI
	Reverse	AAACTGCAG-CTATATGCGCGGCAGGACGCTCAACGG	Pst I
916	Forward	CGCGGATCCCATATG-GCAATGATGGCGGCTG	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTGGCGGCATCTTTCAT	XhoI

917	Forward	AAAAAAGAATTC-CCTGCCGAAAAACCGGCACCGGC	Eco RI
	Reverse	AAAAAACTGCAG-TTATTTCCCCGCCTTCACATCCTG	Pst I
919	Forward	CGCGGATCCCATATG-TGCCAAAGCAAGAGCATC	BamHI-NdeI
	Reverse	CCCGCTCGAG-CGGGCGGTATTCGGG	XhoI
920	Forward	CGCGGATCCCATATG-CACCGCGTCTGGGTC	BamHI-NdeI
	Reverse	CCCGCTCGAG-ATGGTGCGAATGACCGA	XhoI
921	Forward	AAAAAAGAATTC-TTGACGGAAATCCCCGTGAATCC	Eco RI
	Reverse	AAAAAACTGCAG-TCATTTCAAGGGCTGCATCTTCAT	Pst I
922	Forward.2	CGCGGATCCGCTAGC-TGTACGGCGATGGAGGC	BamHI-NheI
	Reverse	CCCGCTCGAG-CAATCCCGGGCCGCC	XhoI
923	Forward	CGCGGATCCCATATG-TGTTACGCAATATTGTCCC	BamHI-NheI
	Reverse	CCCGCTCGAG-GGACAAGGCGACGAAG	XhoI
925	Forward	CGCGGATCCCATATG-AAACAAATGCTTTTAGCCG	BamHI-NdeI
	Reverse	CCCGCTCGAG-GCCGTTGCATTTGATTTC	XhoI
926	Forward	CGCGGATCCCATATG-TGCGCGCAATTACCTC	BamHI-NdeI
	Reverse	CCCGCTCGAG-TCTCGTGCGCGCCG	XhoI
927	Forward	CGCGGATCCCATATG-TGCAGCCCCGCAGC	BamHI-NdeI
	Reverse	CCCGCTCGAG-GTTTTTTGCTGACGTAGT	XhoI
929a	Forward	AAAAAAGAATTC-CGCGGTTTGCTCAAAACAGGGCTGGG	Eco RI
	Reverse	AAAAAATCTAGA-TTAAGAAAGACGGAAACTACTGCC	Xba I
931	Forward	AAAAAAGAATTC-GCAACCCATGTTTTGATGGAAAC	Eco RI
	Reverse	AAAAAACTGCAG-TTACTGCCCGACAACAACGCGACG	Pst I
935	Forward	AAAAAAGAATTC-GCGGATGCGCCCGCGATTTTGGATGACAAGGC	Eco RI
	Reverse	AAAAAACTGCAG-TCAAAACCGCCAATCCGCCGACAC	Pst I
936	Forward	CGCGGATCCCATATG-GCCGCCGTCGGCGC	BamHI-NdeI
	Reverse	CCCGCTCGAG-GCGTTGGACGTAGTTTTG	XhoI
937	Forward	AAAAAAGAATTC-CCGGTTTACATTCAAACCGGCGCAAC	Eco RI
	Reverse	AAAAAACTGCAG-TTAAAATGTATGCTGTACGCCAAA	Pst I
939a	Forward	AAAAAAGAATTC-GGTTCGGCAGCTGTGATGAAACC	Eco RI
	Reverse	AAAAAACTGCAG-TTAACGCAAACCTTGGATAAAGTTGGC	Pst I
950	Forward	CGCGGATCCCATATG-GCCAACAAACCGGCAAG	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTAGAACCGCATTTGCC	XhoI
953	Forward	CGCGGATCCCATATG-GCCACCTACAAAGTGGAC	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTGTTTGGCTGCCTCGAT	XhoI
957	Forward	CGCGGATCCCATATG-TTTTGGCTGGGAACGGG	BamHI-

			NdeI
	Reverse	CCCGCTCGAG-GTGAGAAAGGTCGCGC	XhoI
958	Forward	CGCGGATCCCATATG-GCCGATGCCGTTGCG	BamHI-NdeI
	Reverse	GCCCAAGCTT-GGGTCGTTTGTTGCGTC	HindIII
959	Forward	CGCGGATCCCATATG-CACCACGACGGACACG	BamHI-NdeI
	Reverse	CCCGCTCGAG-GTCGTCGCGGCGGGA	XhoI
961	Forward	CGCGGATCCCATATG-GCCACAAGCGACGACG	BamHI-NdeI
	Reverse	CCCGCTCGAG-CCACTCGTAATTGACGC	XhoI
972	Forward	AAAAAAGAATTC-TTGACTAACAGGGGGGGAGCGAAATTAAAAAC	Eco RI
	Reverse	AAAAAATCTAGA-TTAAAAATAATCATAATCTACATTTTG	Xba I
973	Forward	AAAAAAGAATTC-ATGGACGGCGCACAACCGAAAAC	Eco RI
	Reverse	AAAAAACTGCAG-TTACTTCACGCGGGTCCCCATCAGCGT	Pst I
982	Forward	CGCGGATCCCATATG-GCAGCAAAAGACGTAC	BamHI-NdeI
	Reverse	CCCGCTCGAG-CATCATGCCGCCCATCC	XhoI
983	Forward	CGCGGATCCCATATG-TTAGCTGTTGCAACAACAC	BamHI-NdeI
	Reverse	CCCGCTCGAG-GAACCGGTAGCCTACG	XhoI
987	Forward	CGCGGATCCCATATG-CCCCCACTGGAAGAAC	BamHI-NdeI
	Reverse	CCCGCTCGAG-TAATAAACCTTCTATGGGC	XhoI
988	Forward	CGCGGATCCCATATG-TCTTTAAATTTACGGGAAAAAG	BamHI-NdeI
	Reverse	GCCCAAGCTT-TGATTTGCCTTTCCGTTTT	HindIII
989	Forward	CCGGAATTCTACATATG-GTCCACGCATCCGGCTA	EcoRI-NdeI
	Reverse	CCCGCTCGAG-TTTGAATTTGTAGGTGTATTGC	XhoI
990	Forward. 2	CGCGGATCCGCTAGC-TTCAGAGCTCAGCTT	BamHI-NheI
	Reverse	CCCGCTCGAG-AAACAGCCATTTGAGCGA	XhoI
992	Forward	CGCGGATCCCATATG-GACGCGCCCGCCCG	BamHI-NdeI
	Reverse	CCCGCTCGAG-CCAAATGCCCAACCATTC	XhoI
993	Forward	CGCGGATCCCATATG-GCAATGCTGATTGAAATCA	BamHI-NdeI
	Reverse	CCCGCTCGAG-GAACACATCGCGCCCG	XhoI
996	Forward	CGCGGATCCCATATG-TGCGGCAGAAAATCCGC	BamHI-NdeI
	Reverse	CCCGCTCGAG-TCTAAACCCCTGTTTTCTC	XhoI
997	Forward	CCGGAATTCTAGCTAGC-CGGCACGCCGACGTT	EcoRI-NheI
	Reverse	CCCGCTCGAG-GACGGCATCGCTCAGG	XhoI

Underlined sequences indicate restriction recognition sites.

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 1>:

g001.seq
```
  1 ATGCTGCCGC AGGGGAAGGC GGCGCGGAGG GTGTCGGCGA ACGAGGTGTC
 51 CGGCAGGGCT TGCGCCCGGA TGGTGCTGGT CATCTGCCAG ACGCTGCCGA
101 AACGCGATAC TTTAAACGGC TCGGGTACGC ATACTTTACC GGTTTGGGCG
151 ATTTTGCCGA GGTCGTTGCG CAGCAAATCG ACAATCATCA CGTTTTCGGC
201 GCGGTTTTTC GGGTCGGTTT GTAACTCGGC GGCGCGGCGT TCGTCTTGTC
251 CGTCGCCCAA AATCGGCGCG GTGCCTTTCA TCGGTTCGGT GCTGATGGTG
301 CCGTCTGAAG CGATGTTGAG GAAGAGTTCG GGCGAGAAAC ACAGCGTCCA
351 CGCGGATTGC CCGGCTTCAT CGGGCAGGTG GGACAATACG GCATAG
```

This corresponds to the amino acid sequence <SEQ ID 2; ORF 001.ng>:

g001.pep
```
  1 MLPQGKAARR VSANEVSGRA CARMVLVICQ TLPKRDTLNG SGTHTLPVWA
 51 ILPRSLRSKS TIITFSARFF GSVCNSAARR SSCPSPKIGA VPFIGSVLMV
101 PSEAMLRKSS GEKHSVHADC PASSGRWDNT A*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3>:

m001.seq
```
  1 ATGCTGCCGC AGGGGAAGGC GGCGCGGAGG ATGTCGGCGA ACGAGGTGTG
 51 CGGcAssCTT ss.GCTTGGA yGGTGCTGGT CATCTGCCAA ACGCTGCCGA
101 AACGCGATAC TTTAAACGGT TCGGGTACGC ATACTGTGCC GGTTTGGGCG
151 ATTTTGCCGA GATCGTTACG CAGCAAATCG ACAATCATCA CGTTTTCGGC
201 GCGGTTTTTC GGGTCTGCTT GCAACTCGGC GGCGCGGCGT TCGTCTTGTC
251 CGTCGCCCAA AATCGGCGCG GTGCCTTTCA TCGGTTCGGT GCTGATGGTG
301 CCGTCCGAAC CGATTTTGAG GAAGAGTTCG GGCGAGAAAC ACAGCGTCCA
351 CGCGGATTGC CCCTCCGCAT CGGGCAGGTG GGACAAGACG GCATAG
```

This corresponds to the amino acid sequence <SEQ ID 4; ORF 001>:

m001.pep
```
  1 MLPQGKAARR MSANEVCGXL XAWXVLVICQ TLPKRDTLNG SGTHTVPVWA
 51 ILPRSLRSKS TIITFSARFF GSACNSAARR SSCPSPKIGA VPFIGSVLMV
101 PSEPILRKSS GEKHSVHADC PSASGRWDKT A*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 5>:

a001.seq
```
  1 ATGCTGCCGC AGGGGAAGGC GGCGCGGAGG ATGTCGGCGA ACGAGGTGTG
 51 CGGCAAGGCT TGGGCTTGGA TGGTGCTGGT CATCTGCCAA ACGCTGCCGA
101 AACGCGATAC TTTAAACGGT TCGGGTACGC ATACTGTGCC GGTTTGGGCG
151 ATTTTGCCGA GGTCGTTACG CAGCAAATCG ACAATCATCA CGTTTTCGGC
201 GCGGTTTTTC GGGTCTGCTT GCAACTCGGC GGCGCGGCGT TCGTCTTGTC
251 CGTCGCCCAA AATCGGCGCG GTGCCTTTCA TCGGTTCGGT GCTGATGGTG
301 CCGTCCGAAC CGATTTTGAG GAAGAGTTCG GGCGAGAAAC ACAGCGTCCA
351 CGCGGATTGC CCTTGTGCAT CGGGCAGGTG GGACAAAACG GCATAG
```

This corresponds to the amino acid sequence <SEQ ID 6; ORF 001.a>:

a001.pep
```
  1 MLPQGKAARR MSANEVCGKA WAWMVLVICQ TLPKRDTLNG SGTHTVPVWA
 51 ILPRSLRSKS TIITFSARFF GSACNSAARR SSCPSPKIGA VPFIGSVLMV
101 PSEPILRKSS GEKHSVHADC PCASGRWDKT A*
```

m001/a001 96.2% identity over a 131 aa overlap

```
                   10        20        30        40        50        60
m001.pep   MLPQGKAARRMSANEVCGXLXAWXVLVICQTLPKRDTLNGSGTHTVPVWAILPRSLRSKS
           ||||||||||||||||||   || ||||||||||||||||||||||||||||||||||||
a001.pep   MLPQGKAARRMSANEVCGKAWAWMVLVICQTLPKRDTLNGSGTHTVPVWAILPRSLRSKS
```

```
                 10         20         30         40         50         60

                 70         80         90        100        110        120
m001.pep  TIITFSARFFGSACNSAARRSSCPSPKIGAVPFIGSVLMVPSEPILRKSSGEKHSVHADC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a001.pep  TIITFSARFFGSACNSAARRSSCPSPKIGAVPFIGSVLMVPSEPILRKSSGEKHSVHADC
                 70         80         90        100        110        120

                130
m001.pep  PSASGRWDKTAX
          | ||||||||||
a001.pep  PCASGRWDKTAX
                130
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 001 shows 89.3% identity over a 131 aa overlap with predicted ORF (ORF 001.ng) from *N. gonorrhoeae:*

m001/g001

```
                 10         20         30         40         50         60
m001.pep  MLPQGKAARRMSANEVCGXLXAWXVLVICQTLPKRDTLNGSGTHTVPVWAILPRSLRSKS
          ||||||||||:||||| |   |  |||||||||||||||||||:||||||||||||||||
g001      MLPQGKAARRVSANEVSGRACARMVLVICQTLPKRDTLNGSGTHTLPVWAILPRSLRSKS
                 10         20         30         40         50         60

                 70         80         90        100        110        120
m001.pep  TIITFSARFFGSACNSAARRSSCPSPKIGAVPFIGSVLMVPSEPILRKSSGEKHSVHADC
          ||||||||||||:|||||||||||||||||||||||||||||| :|||||||||||||||
g001      TIITFSARFFGSVCNSAARRSSCPSPKIGAVPFIGSVLMVPSEAMLRKSSGEKHSVHADC
                 70         80         90        100        110        120

                130
m001.pep  PSASGRWDKTAX
          |::|||||:|||
g001      PASSGRWDNTAX
                130
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 7>:

g003.seq

```
  1  ATGGTCGTAT TCGTGGCTGA AGGCGTATTC GGTCGCGCTG TTTTGGGTCA
 51  CTTGGTATTG CTCTTCGGTC AGGGTGCGTT TGAGTTCGGC GTCACTCGGT
101  TTTTTATACG TTGCCGCGTC GAAGCCTTTG CCTTGCGGTG CGGCTTTGGT
151  TTTGCCCGGC AGCGGTTCGT CGGCTTTGCG GATGTCGATG TGGCAGTAGC
201  CGTTGGGGTT TTTAATCAGG TAGTCCTGAT GGTATTCCTC GGCGTCGTAG
251  AAGTTTTTCA GCGGTTCGTT TTCAACAACG AGGGGCAGTT GGTATTTTTG
301  CTGCTCGCGT TTGAGGGCGG CGGCGATGAC GGCTTTTTCG GCGGGGTCGG
351  TGTAGTACAC GCCGCTGCGG TATTGCGTGC CGGTGTCGTT ACCCTGTTTG
401  TTGAGGCTGG TCGGATCAAC GACGCGGAAA TAATATTGCA GGATGTCGTC
451  CAGgCTGagt TTGTCGGCAT CGTaggtcac tTTGACGGTC TCGGCATGAC
501  CCGTATGGCG GTaggacact tctTCgtanc TcGGGtTTTC CGTGttGCCG
551  TTGGCgttac cGGATACCGC gtcaACCACG CCGTcgatgc gttggaAATa
601  ggCTTCCAAg ccccaaaagc agccgccggc gaagtaaatg gtgcccgtgt
651  tcatgattGC TGa
```

This corresponds to the amino acid sequence <SEQ ID 8; ORF 003.ng>:

g003.pep

```
  1  MVVFVAEGVF GRAVLGHLVL LFGQGAFEFG VTRFFIRCRV EAFALRCGFG
 51  FARQRFVGFA DVDVAVAVGV FNQVVLMVFL GVVEVFQRFV FNNEGQLVFL
101  LLAFEGGGDD GFFGGVGVVH AAAVLRAGVV TLFVEAGRIN DAEIILQDVV
```

```
151  QAEFVGIVGH FDGLGMTRMA VGHFFVRVFR VAVGVTGYRV NHAVDALEIG
201  FQAPKAAAGE VNGARVHDC
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 9>:

m003.seq

```
  1  ATGGTCGTAT TCGTGGCTGA AGGCATATTC GGTCGCGCTG TTTTGGGTAA
 51  CTTGGsTATTG CTCTTCGGTC AGGGTGCGTT TGAGTTCGGC GTCACTCGGT
101  TTTTTATACG TTGCCGCGTC GAAGCCTTTG CCTTGCGGGG CGGTCTTGGT
151  TTTGCCCGGC AGCGGTTCGT CAGCkTTGCG GATGTCGATG TGGCAGTAGC
201  CGTTGGGGTT TTTAATCAAG TAGTCCTGAT GGTATTCCTC GGCATCGTAG
251  AAGTTTTtCA GCGGCTCGTT TTCAACAACG AGGGGCAGTT GGTATTTTTG
301  CTGCTCGCGT TTGAGGGCGk CGGCGATGAC GGCTTTTTCG kCGGGGTCGG
351  TGTAGTACAC GCCGCTGCGG TATTGCGTAC CGGTGTCGTT GCCCTGTTTG
401  TTGAGGCTGG TCGGATCAAC GACGCGGAAG AAATATTGCA GGATGTCGTC
451  TAGGCTGAGT TTGTCGGCAT CGTAGGTCAC TTTGACGGTT TCGGCGTGGC
501  CCGTATGGCG GTAGGACACG TCTTCATAGC TCGGATTTTT CGTGTTGCCG
551  TTGGCGTAGC CGGATACCGC GTCAACCACG CCGTCGATGC GTTGGAAATA
601  GGCTTCCAAG CCCCAGAAGC AGCg.CCGGC GAGGTAAATG GTGCGCGTGT
651  TCATGATTTT TGA
```

This corresponds to the amino acid sequence <SEQ ID 10; ORF 003>:

m003.pep Length: 221

```
  1  MVVFVAEGIF GRAVLGNLXL LFGQGAFEFG VTRFFIRCRV EAFALRGGLG
 51  FARQRFVSXA DVDVAVAVGV FNQVVLMVFL GIVEVFQRLV FNNEGQLVFL
101  LLAFEGXGDD GFFXGVGVVH AAAVLRTGVV ALFVEAGRIN DAEEILQDVV
151  *AEFVGIVGH FDGFGVARMA VGHVFIARIF RVAVGVAGYR VNHAVDALEI
201  GFQAPEAAXG EVNGARVHDF *
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 11>:

a003.seq

```
  1  ATGGTCGTAT TCGTGGCTGA AGGCATATTC GGTCGCGCTG TTTTGGGTAA
 51  CTTGGTATTG CTCTTCGGTC AGGGTGCGTT TGAGTTCGGC GTCACTCGGT
101  TTTTTATACG TTGCCGCGTC GAAGCCTTTG CCTTGCGGTG CGGTCTTGGT
151  TTTGCCCGGC AGCGGTTCGT CGGCTTTGCG GATATCGATG TGGCAGTAGC
201  CGTTGGGGTT TTTAATCAAG TAGTCCTGAT GGTATTCCTC GGCATCGTAG
251  AAGTTTTTCA GCGGCTCGTT TTCAACAACG AGGGGCAGTT GGTATTTTTG
301  CTGCTCGCGT TTGAGGGCGG CGGCGATGAC GGCTTTTTCG GCGGGGTCGG
351  TGTAGTACAC GCCGCTGCGG TATTGCGTAC CGGTGTCGTT GCCCTGTTTG
401  TTGAGGCTGG TCGGATCAAC GACGCGGAAG AAATATTGCA GGATGTCGTC
451  TAGGCTGAGT TTGTCGGCAT CGTAGGTCAC TTTGACGGTT TCGGCGTGGC
501  CCGTATGGCG GTAGGACACG TCTTCATAGC TCGGATTTTT CGTGTTGCCG
551  TTGGCGTAGC CGGATACCGC GTCAACCACG CCGTCGATGC GTTGGAAATA
601  GGCTTCCAAG CCCCAGAAGC AGCCGCCGGC GAGGTAGATG GTGCGCGTGT
651  TCATGATTTT TGA
```

This corresponds to the amino acid sequence <SEQ ID 12; ORF 003.a>:

a003.pep

```
  1  MVVFVAEGIF GRAVLGNLVL LFGQGAFEFG VTRFFIRCRV EAFALRCGLG
 51  FARQRFVGFA DIDVAVAVGV FNQVVLMVFL GIVEVFQRLV FNNEGQLVFL
101  LLAFEGGGDD GFFGGVGVVH AAAVLRTGVV ALFVEAGRIN DAEEILQDVV
151  *AEFVGIVGH FDGFGVARMA VGHVFIARIF RVAVGVAGYR VNHAVDALEI
201  GFQAPEAAAG EVDGARVHDF *
```

m003/a003 95.9% identity over a 220 aa overlap

```
                    10        20        30        40        50        60
m003.pep    MVVFVAEGIFGRAVLGNLXLLFGQGAFEFGVTRFFIRCRVEAFALRGGLGFARQRFVSXA
            ||||||||||||||||||| ||||||||||||||||||||||||||| ||||||||||: |
a003        MVVFVAEGIFGRAVLGNLVLLFGQGAFEFGVTRFFIRCRVEAFALRCGLGFARQRFVGFA
                    10        20        30        40        50        60
```

```
                      70        80        90       100       110       120
m003.pep      DVDVAVAVGVFNQVVLMVFLGIVEVFQRLVFNNEGQLVFLLLAFEGXGDDGFFXGVGVVH
              |:|||||||||||||||||||||||||||||||||||||||||||| |||||| ||||||
a003          DIDVAVAVGVFNQVVLMVFLGIVEVFQRLVFNNEGQLVFLLLAFEGGGDDGFFGGVGVVH
                      70        80        90       100       110       120

                     130       140       150       160       170       180
m003.pep      AAAVLRTGVVALFVEAGRINDAEEILQDVVXAEFVGIVGHFDGFGVARMAVGHVFIARIF
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a003          AAAVLRTGVVALFVEAGRINDAEEILQDVVXAEFVGIVGHFDGFGVARMAVGHVFIARIF
                     130       140       150       160       170       180

                     190       200       210       220
m003.pep      RVAVGVAGYRVNHAVDALEIGFQAPEAAXGEVNGARVHDFX
              |||||||||||||||||||||||||||| |||:||||||||
a003          RVAVGVAGYRVNHAVDALEIGFQAPEAAAGEVDGARVHDFX
                     190       200       210       220
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 003 shows 88.6% identity over a 219 aa overlap with a predicted ORF (ORF 003.ng) from *N. gonorrhoeae:*

m003/g003

```
                      10        20        30        40        50        60
m003.pep      MVVFVAEGIFGRAVLGNLXLLFGQGAFEFGVTRFFIRCRVEAFALRGGLGFARQRFVSXA
              ||||||||:|||||||:| |||||||||||||||||||||||||||| :||||||||: |
g003          MVVFVAEGVFGRAVLGHLVLLFGQGAFEFGVTRFFIRCRVEAFALRCGFGFARQRFVGFA
                      10        20        30        40        50        60

                      70        80        90       100       110       120
m003.pep      DVDVAVAVGVFNQVVLMVFLGIVEVFQRLVFNNEGQLVFLLLAFEGXGDDGFFXGVGVVH
              |||||||||||||||||||||:||||||:||||||||||||||||| |||||| ||||||
g003          DVDVAVAVGVFNQVVLMVFLGVVEVFQRFVFNNEGQLVFLLLAFEGGGDDGFFGGVGVVH
                      70        80        90       100       110       120

                     130       140       150       160       170       180
m003.pep      AAAVLRTGVVALFVEAGRINDAEEILQDVVXAEFVGIVGHFDGFGVARMAVGHVFIARIF
              ||||||:|||:|||||||||||| |||||| ||||||||||||:|::|||||| |: |:|
g003          AAAVLRAGVVTLFVEAGRINDAEIILQDVVQAEFVGIVGHFDGLGMTRMAVGHFFV-RVF
                     130       140       150       160       170       180

                     190       200       210       220
m003.pep      RVAVGVAGYRVNHAVDALEIGFQAPEAAXGEVNGARVHDFX
              ||||||:||||||||||||||||||:|| ||||||||||
g003          RVAVGVTGYRVNHAVDALEIGFQAPKAAAGEVNGARVHDC
                     190       200       210
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 13>:

g004.seq

```
  1   ATGgtagAAC GGCATATCCA GCATTTGCGG AACGGTCATC TTCATTTGAT
 51   GCGCCCATGC CAACAagtga gccaAAtgtT CGGCGGCAGG GCCTacgatT
101   TCCGCGCCGA TAAagcggcc gGTGgctTTT tcgGCataca ggcgcaTatg
151   gCCTTTGTTT ACCAgcatca cgcggctgcg accttgaTTT TTGAACGATA
201   CTTCGCCgaT GACAAATTCG TCGGCTTGGT ATTGCGCGGC AACCTGCGCG
251   TATTTCAAAC CGACAAAGCC GATTTGCgga ctggtaaACA CCACGCCAAT
301   GGTgctgcgg cGCAAACCGC TGCCGATATt cgGgtagcgg ccccgcgtta
351   ttgcccggca atcttacctt ggtcggcggc ttcatGCAGC AGGGGCagtt
401   ggttggacgc gtcgcccgca ataAAGATAT GCGGAATgct ggtCTGCATg
451   gtCAGCGGAT CGGCAACGGG tacgccgcgc gcgtctttgT CGATATTGAT
501   GTTTTCCAAA CCGATATTtgT CAACGTTCGG ACGGCgACCT ACGGCTGCCA
```

```
551  ACATATATTC GGCAACAAA. ACGCCTTTTT CGCCATCCTG CTCCCAATGG
601  ACTtctACAT TGCCGTCTGC GTCGAGTTTG ACCTCGGTTT TAGCATCCAG
651  ATGCAGTTTC AATtctTCTC CGAACACGGC TTTCGCCTCG TCTGAAACAA
701  CGGGGTCGGA AATGCCGCCG ATGATTCCGC CCAAACCGAA AATTTCAACT
751  TTCACACCCA AACGGTGCAA TGCCTGA
```

This corresponds to the amino acid sequence <SEQ ID 14; ORF 004.ng>:

g004.pep

```
  1  MVERHIQHLR NGHLHLMRPC QQVSQMFGGR AYDFRADKAA GGFFGIQAHM
 51  AFVYQHHAAA TLIFERYFAD DKFVGLVLRG NLRVFQTDKA DLRTGKHHAN
101  GAAAQTAADI RVAAPRYCPA ILPWSAASCS RGSWLDASPA IKICGMLVCM
151  VSGSATGTPR ASLSILMFSK PILSTFGRRP TAANIYSATN TPFSPSCSQW
201  TSTLPSASSL TSVLASRCSF NSSPNTAFAS SETTGSEMPP MIPPKPKIST
251  FTPKRCNA*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 15>:

m004.seq

```
  1  ATGGTAGAAC GGCATATCCA GCATTTGCGG AACGGTCATC TTCATTTGAT
 51  GTGCCCAAGC CAACAGGTGC GCCAAATGTT CGGCGGCAGG GCCTACGATT
101  TCCGCGCCGA TAAAGCGGCC GGTGGCTTTT TCGGCATACA GGCGCATATG
151  GCCTTTGTTC ACCAGCATCA CGCGGCTGCG GCCTTGGTTT TTGAACGATA
201  CTTCGCCGAT GACAAATTCG TCGGCTTGGT ATTGCGCGGC AACCTGCGCG
251  TATTTCAGAC CGACAAAGCC GATTTGCGGA CTGGTAAACA CCACGCCGAT
301  GGTGCTGCGC CGCAAACCGC CGCCGATATT CGGGTAGCGG CCGCGTTATC
351  GCCGGCAATC TTGCCTTGGT CGGCAGCTTC ATGCAGCAGA GGCAGTTGGT
401  TGGACGCATC GCCTGCGATG AAGATATGCG GAATACTGGT CTGCATGGTC
451  AGCGGGTCGG CAACAGGTAC GCCGCGCGCA TCTTTTTCGA TATTGATATT
501  TTCCAAACCG ATATTGTCAA CGTTCGGACG GCGGCCCACG GCTGCCAGCA
551  TATATTCGGC AACAAATACG CCTTTTTCGC CATCCTGCTC CCAATGGACT
601  TCTACATTGC CGTCTGCATC GAGTTTGACC TCGGTTTTAG CATCCAGATG
651  CAGTTTCAAT TCTTCGCCGA ACACGGCGTT CGCCTCGTCT GAAACGACGG
701  GGTCGGAAAT GCCGCCGATG ATTCCGCCCA AACCGAAAAT TTCAACTTTC
751  ACGCCCAAAC GGTGCAATGC CTGA
```

This corresponds to the amino acid sequence <SEQ ID 16; ORF 004>:

m004.pep

```
  1  MVERHIQHLR NGHLHLMCPS QQVRQMFGGR AYDFRADKAA GGFFGIQAHM
 51  AFVHQHHAAA ALVFERYFAD DKFVGLVLRG NLRVFQTDKA DLRTGKHHAD
101  GAAPQTAADI RVAAALSPAI LPWSAASCSR GSWLDASPAM KICGILVCMV
151  SGSATGTPRA SFSILIFSKP ILSTFGRRPT AASIYSATNT PFSPSCSQWT
201  STLPSASSLT SVLASRCSFN SSPNTAFASS ETTGSEMPPM IPPKPKISTF
251  TPKRCNA*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 17>:

a004.seq

```
  1  ATGGTAGAAC GGCATATCCA GCATTTGCGG AACGGTCATC TTCATTTGAT
 51  GTGCCCAAGC CAACAGGTGC GCCAAATGTT CGGCGGCCGG ACCTACGATT
101  TCTGCGCCGA TGAAGCGGCC GGTGGCTTTT TCGGCATACA GGCGCATATG
151  GCCTTTGTTT ACCAGCATCA CGCGGCTGCG GCCTTGGTTT TTGAACGATA
201  CTTCGCCGAT GACAAATTCG TCGGCTTGGT ATTGCGCGGC AACCTGCGCG
251  TATTTCAAAC CGACAAAGCC GATTTGCGGA CTGGTGAACA CTACGCCGAT
301  GGTGCTGCGG CGCAAACCGC CGCCGATATT CGGGTAGCGG CCGCGTTATC
351  GCCGGCAATC TTGCCTTGGT CGGCGGCTTC ATGCAGCAGG GGCAGTTGGT
401  TGGACGCGTC GCCCGCAATA AAGATATGCG GAATACTGGT CTGCATAGTC
451  AGCGGATCGG CAACGGGTAC GCCGCGCGCA TCTTTTTCGA TATTGATGTT
501  TTCCAAACCG ATATTGTCAA CGTTCGGACG GCGGCCTACG GCTGCCAGCA
551  TATATTCGGC AACAAATACG CCTTTTTCGC CATCCTGCTC CCAATGGACT
601  TCTACATTGC CGTCTGCGTC GAGTTTGGCC TCGGTTTTAG CATCCAAATG
651  CAGTTTCAAT TCTTCACCGA ACACGGCTTT CGCCTCGTCT GAAACGACGG
701  GGTCGGAAAT GCCGCCGATG ATGCCACCCA AACCGAAAAT TTCAACTTTC
751  ACGCCCAAAC GGTGCAATGC CTGA
```

This corresponds to the amino acid sequence <SEQ ID 18; ORF 004.a>:

a004.pep

```
  1 MVERHIQHLR NGHLHLMCPS QQVRQMFGGR TYDFCADEAA GGFFGIQAHM
 51 AFVYQHHAAA ALVFERYFAD DKFVGLVLRG NLRVFQTDKA DLRTGEHYAD
101 GAAAQTAADI RVAAALSPAI LPWSAASCSR GSWLDASPAI KICGILVCIV
151 SGSATGTPRA SFSILMFSKP ILSTFGRRPT AASIYSATNT PFSPSCSQWT
201 STLPSASSLA SVLASKCSFN SSPNTAFASS ETTGSEMPPM MPPKPKISTF
251 TPKRCNA*
```

m004/a004 94.9% identity over a 257 aa overlap

```
                    10        20        30        40        50        60
m004.pep    MVERHIQHLRNGHLHLMCPSQQVRQMFGGRAYDFRADKAAGGFFGIQAHMAFVHQHHAAA
            ||||||||||||||||||||||||||||||:||| ||:|||||||||||||||:||||||
a004        MVERHIQHLRNGHLHLMCPSQQVRQMFGGRTYDFCADEAAGGFFGIQAHMAFVYQHHAAA
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m004.pep    ALVFERYFADDKFVGLVLRGNLRVFQTDKADLRTGKHHADGAAPQTAADIRVAAALSPAI
            |||||||||||||||||||||||||||||||||||:|:||||| ||||||||||||||||
a004        ALVFERYFADDKFVGLVLRGNLRVFQTDKADLRTGEHYADGAAAQTAADIRVAAALSPAI
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m004.pep    LPWSAASCSRGSWLDASPAMKICGILVCMVSGSATGTPRASFSILIFSKPILSTFGRRPT
            |||||||||||||||||||:||||||||:||||||||||||||||:||||||||||||||
a004        LPWSAASCSRGSWLDASPAIKICGILVCIVSGSATGTPRASFSILMFSKPILSTFGRRPT
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m004.pep    AASIYSATNTPFSPSCSQWTSTLPSASSLTSVLASRCSFNSSPNTAFASSETTGSEMPPM
            |||||||||||||||||||||||||||||:|||||:||||||||||||||||||||||||
a004        AASIYSATNTPFSPSCSQWTSTLPSASSLASVLASKCSFNSSPNTAFASSETTGSEMPPM
                   190       200       210       220       230       240

                   250
m004.pep    IPPKPKISTFTPKRCNAX
            :|||||||||||||||||
a004        MPPKPKISTFTPKRCNAX
                   250
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

ORF 004 shows 93.4% identity over a 258 aa overlap with a predicted ORF (ORF 004.ng) from *N. gonorrhoeae:*

m004/g004

```
                    10        20        30        40        50        60
m004.pep    MVERHIQHLRNGHLHLMCPSQQVRQMFGGRAYDFRADKAAGGFFGIQAHMAFVHQHHAAA
            ||||||||||||||||| | ||| |||||||||||||||||||||||||||||:||||||
g004        MVERHIQHLRNGHLHLMRPCQQVSQMFGGRAYDFRADKAAGGFFGIQAHMAFVYQHHAAA
                    10        20        30        40        50        60

                    70        80        90       100       110       119
m004.pep    ALVFERYFADDKFVGLVLRGNLRVFQTDKADLRTGKHHADGAAPQTAADIRVAAA-LSPA
            :|:||||||||||||||||||||||||||||||||||||:||| |||||||||||   ||
g004        TLIFERYFADDKFVGLVLRGNLRVFQTDKADLRTGKHHANGAAAQTAADIRVAAPRYCPA
                    70        80        90       100       110       120

          120       130       140       150       160       170       179
m004.pep    ILPWSAASCSRGSWLDASPAMKICGILVCMVSGSATGTPRASFSILIFSKPILSTFGRRP
```

```
             ||||||||||||||||||||:||||:||||||||||||||||:|||:||||||||||||
g004         ILPWSAASCSRGSWLDASPAIKICGMLVCMVSGSATGTPRASLSILMFSKPILSTFGRRP
                     130       140       150       160       170       180

          180       190       200       210       220       230       239
m004.pep     TAASIYSATNTPFSPSCSQWTSTLPSASSLTSVLASRCSFNSSPNTAFASSETTGSEMPP
             |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g004         TAANIYSATNTPFSPSCSQWTSTLPSASSLTSVLASRCSFNSSPNTAFASSETTGSEMPP
                     190       200       210       220       230       240

          240       250
m004.pep     MIPPKPKISTFTPKRCNAX
             ||||||||||||||||||
g004         MIPPKPKISTFTPKRCNA
                     250
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 19>:

g005.seq

```
   1  ATGGGGATGG ACAATATTGA TATGTTCATG CCTGAACAAG AGGAAATCCA
  51  ATCAATGTGG AAAGAAATTT TACTGAATTA CGGTATTTTC CTGCTCGAAC
 101  TGCTTACCGT GTTCGGCGCA ATTGCGCTGA TTGTGTTGGC TATCGTACAG
 151  AGTAAGAAAC AGTCGGAAAG CGGCAGTGTC GTACTGACAG ATTTTTCGGA
 201  AAATTATAAA AAACAGCGGC AATCGTTTGA AACATTCTTT TTAAGCGAGG
 251  AAGAGACAAA ACATCAGGAA AAAAAAGAAA AGAAAAAGGA AAAGGCGGAA
 301  GCCAAAGCAG AGAAAAAGCG TTTGAAGGAG GGCGGGGAGA AATCTGCCGA
 351  AACGCAAAAA TCCCGCCTTT TTGTGTTGGA TTTTGACGGC GATTTGTATG
 401  CACACGCCGT AGAATCCTTG CGTCATGAGA TTACGGCGGT GCTTTTGATT
 451  GCCAAGCCTG AAGATGAGGT TCTGCTCAGA TTGGAAAGTC CGGGCGGCGT
 501  GGTTCACGGT TACGGTTTGG CGGCTTCGCA GCTTAGGCGT TTGCGCGAAC
 551  GCAATATTCC GCTGAccgtc gccgTCGATA AGGTCGCGGC AAGCGgcggc
 601  tatatgatgg cgtgtgtgGC GGATAAAATT GTTTCCGCtc cgtttgcggt
 651  catcggttcg gtgggtgtgg tgGcggaagt gcCGAATATC CAccgCctGT
 701  TGAAAAAACA TGATATTGAT GTGGATGTGA TGACGGCGGG CGAATTTAAG
 751  CGCACGGTTA CTTTTATGGG TGAAAATACG GAAAAGGGCA AACAGAAATT
 801  CCGGCAGGAA CTGGAGGAAA CGCATCAGTT GTTCAAGCAG TTTGTCAGTG
 851  AAAACCGCCC CGGGTTGGAT ATTGAAAAAA TAGCGACGGG CGAGCATTGG
 901  TTCGGCCGGC AGGCGTTGGC GTTGAACTTG ATTGACGAGA TTTCGACCAG
 951  TGATGATTTG TTGTTGAAAG CGTTTGAAAA CAAACAGGtt aTCGAAGTGA
1001  AATATCAGGA GAAGCGAAGC CTGATCCAGC GCATTGGTTT GCAGGCGGAA
1051  GCTTCCGTTG AAAAGTTGTT TGCCAAACTT GTCAACCGGC GAGCGGATGT
1101  GATGTAG
```

This corresponds to the amino acid sequence <SEQ ID 20; ORF 005.ng>:

g005.pep

```
   1  MGMDNIDMFM PEQEEIQSMW KEILLNYGIF LLELLTVFGA IALIVLAIVQ
  51  SKKQSESGSV VLTDFSENYK KQRQSFETFF LSEEETKHQE KKEKKKEKAE
 101  AKAEKKRLKE GGEKSAETQK SRLFVLDFDG DLYAHAVESL RHEITAVLLI
 151  AKPEDEVLLR LESPGGVVHG YGLAASQLRR LRERNIPLTV AVDKVAASGG
 201  YMMACVADKI VSAPFAVIGS VGVVAEVPNI HRLLKKHDID VDVMTAGEFK
 251  RTVTFMGENT EKGKQKFRQE LEETHQLFKQ FVSENRPGLD IEKIATGEHW
 301  FGRQALALNL IDEISTSDDL LLKAFENKQV IEVKYQEKRS LIQRIGLQAE
 351  ASVEKLFAKL VNRRADVM*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 21>:

m005.seq

```
   1  ATGGACAATA TTGACATGTT CATGCCTGAA CAAGAGGAAA TCCAATCAAT
  51  GTGGAAAGAA ATTTTACTGA ATTACGGTAT TTTCCTGCTC GAACTGCTTA
 101  CCGTGTTCGG CGCAATTGCG CTGATTGTGT TGGCTATCGT ACAGAGTAAG
 151  AAACAGTCGG AwAGCGGCAG TGTCGTACTG ACGGATTTTT CGGAAAATTA
 201  TAAAAAACAG CGGCAATCGT TTGAAGCATT CTTTTTAAGC GGGGAAGAGG
 251  CACAACATCA GGAAAAAGAG GAAAAGAAAA AGGAAAAGGC GGAAGCCAAA
```

```
 301  GCAGAGAAAA A.CGTTTGAA GGAGGGTGGG GAGAAATCTG CCGAAACGCA
 351  nAAATCACGC CTTTTTGTGT TGGANNNNNN NNNNNNNNNN NNNNNNNNNN
 401  NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 451  NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 501  NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 551  NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNGCGAGCGG CGGTTATATG
 601  ATGGCGTGTG TGGCGGATAA AATTGCTTCC GCTCCGTTTG CGATTGTCGG
 651  TTCGGTGGGT GTGGTGGCGG AAGTACCGAA TATCCACCGC CTGTTGAAAA
 701  AACATGATAT TGATGTGGAT GTGATGACGG CGGGCGAATT TAAGCGCACG
 751  GTTACTTTTA TGGGTGAAAA TACGGAAAAG GGCAAACAGA AATTCCGACA
 801  GGAACTGGAG GAAACGCATC AGTTGTTCAA GCAGTTTGTC AGCGAGAACC
 851  GCCCTCAATT GGATATTGAG GAAGTGGCAA CGGGCGAGCA TTGGTTCGGT
 901  CGGCAGGCGT TGGCGTTGAA CTTGATTGAC GAGATTTCGA CCAGTGATGA
 951  TTTGTTGTTG AAAGCGTTTG AAAACAAACA GGTTATCGAA GTGAAATATC
1001  AGGAGAAGCA AAGCCTGATC CAGCGCATTG GTTTGCAGGC GGAAGCTTCT
1051  GTTGAAAAGT TGTTTGCCAA ACTTGTCAAC CGGCGGGCGG ATGTGATGT A
1101  G
```

This corresponds to the amino acid sequence <SEQ ID 22; ORF 005>:

m005.pep

```
   1  MDNIDMFMPE QEEIQSMWKE ILLNYGIFLL ELLTVFGAIA LIVLAIVQSK
  51  KQSXSGSVVL TDFSENYKKQ RQSFEAFFLS GEEA[illegible]KE EKKKEKAEAK
 101  AEKXRLKEGG EKSAETXKSR LFVLXXXXXX XXXXXXXXXX XXXXXXXXXX
 151  XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXASGGYM
 201  MACVADKIAS APFAIVGSVG VVAEVPNIHR LLKKHDIDVD VMTAGEFKRT
 251  VTFMGENTEK GKQKFRQELE ETHQLFKQFV SENRPQLDIE EVATGEHWFG
 301  RQALALNLID EISTSDDLLL KAFENKQVIE VKYQEKQSLI QRIGLQAEAS
 351  VEKLFAKLVN RRADVM*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 23>:

a005.seq

```
   1  ATGGACAATA TTGACATGTT CATGCCTGAA CAAGAGGAAA TCCAATCAAT
  51  GTGGAAAGAA ATTTTACTGA ATTACGGTAT TTTCCTGCTC GAACTGCTTA
 101  CCGTGTTCGG CGCAATTGCG CTGATTGTGT TGGCTATCGT ACAGAGTAAG
 151  AAACAGTCGG AAAGCGGCAG TGTCGTACTG ACGGATTTTT CGGAAAATTA
 201  TAAAAAACAG CGGCAATCGT TTGAAGCATT CTTTTTAAGC GGGGAAGAGG
 251  CAAAACATCA GGAAAAAGAG GAAAAGAAAA AGGAAAAGGC GGAAGCCAAA
 301  GCAGAGAAAA AGCGTTTGAA GGAGGGTGGG GAGAAATCTT CCGAAACGCA
 351  AAAATCCCGC CTTTTTGTGT TGGATTTTGA CGGCGATTTG TATGCACACG
 401  CCGTAGAATC CTTGCGTCAT GAGATTACGG CGGTGCTTTT GATTGCCAAG
 451  CCTGAAGATG AGGTTCTGCT TAGATTGGAA AGTCCGGGCG GCGTGGTTCA
 501  CGGTTACGGT TTGGCGGCTT CGCAGCTTAG GCGTTTGCGC GAACGCAATA
 551  TTCCGCTGAC CGTCGCCGTC GATAAGGTGG CGGCGAGCGG TGGTTATATG
 601  ATGGCGTGTG TGGCGGATAA AATTGTTTCC GCTCCGTTTG CGATTGTCGG
 651  TTCGGTGGGT GTTGTAGCGG AAGTACCGAA TATCCACCGC CTGTTGAAAA
 701  AACATGATAT TGATGTGGAT GTGATGACGG CGGGCGAATT TAAGCGCACG
 751  GTTACTTTTA TGGGTGAAAA TACGGAAAAG GGCAAACAGA AATTCCGACA
 801  GGAACTGGAG GAAACGCATC AGTTGTTCAA GCAGTTTGTC AGCGAGAACC
 851  GCCCTCAATT GGATATTGAG GAAGTGGCAA CGGGCGAGCA TTGGTTCGGT
 901  CGGCAGGCGT TGGCGTTGAA CTTGATTGAC GAGATTTCGA CCAGTGATGA
 951  TTTGTTGTTG AAAGCGTTTG AAAACAAACA GGTTATCGAA GTGAAATATC
1001  AGGAGAAGCA AAGCCTGATC CAGCGCATTG GTTTGCAGGC GGAAGCTTCT
1051  GTTGAAAAGT TGTTTGCCAA ACTTGTCAAC CGGCGGGCGG ATGTGATGTA
1101  G
```

This corresponds to the amino acid sequence <SEQ ID 24; ORF 005.a>:

a005.pep

```
   1  MDNIDMFMPE QEEIQSMWKE ILLNYGIFLL ELLTVFGAIA LIVLAIVQSK
  51  KQSESGSVVL TDFSENYKKQ RQSFEAFFLS GEEAKHQEKE EKKKEKAEAK
 101  AEKKRLKEGG EKSSETQKSR LFVLDFDGDL YAHAVESLRH EITAVLLIAK
 151  PEDEVLLRLE SPGGVVHGYG LAASQLRRLR ERNIPLTVAV DKVAASGGYM
 201  MACVADKIVS APFAIVGSVG VVAEVPNIHR LLKKHDIDVD VMTAGEFKRT
 251  VTFMGENTEK GKQKFRQELE ETHQLFKQFV SENRPQLDIE EVATGEHWFG
```

```
301  RQALALNLID EISTSDDL__ KAFENKQVIE VKYQEKQSLI QRIGLQAEAS
351  VEKLFAKLVN RRADVM*
```

m005/a005 79.2% identity over a 366 aa overlap

```
                   10        20        30        40        50        60
m005.pep   MDNIDMFMPEQEEIQSMWKEILLNYGIFLLELLTVFGAIALIVLAIVQSKKQSXSGSVVL
           |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
a005       MDNIDMFMPEQEEIQSMWKEILLNYGIFLLELLTVFGAIALIVLAIVQSKKQSESGSVVL
                   10        20        30        40        50        60

                   70        80        90       100       110       120
m005.pep   TDFSENYKKQRQSFEAFFLSGEEAQHQEKEEKKKEKAEAKAEKXRLKEGGEKSAETXKSR
           ||||||||||||||||||||||||:||||||||||||||||||| ||||||||:|| |||
a005       TDFSENYKKQRQSFEAFFLSGEEAKHQEKEEKKKEKAEAKAEKKRLKEGGEKSSETQKSR
                   70        80        90       100       110       120

                  130       140       150       160       170       180
m005.pep   LFVLXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
           ||||                    :
a005       LFVLDFDGDLYAHAVESLRHEITAVLLIAKPEDEVLLRLESPGGVVHGYGLAASQLRRLR
                  130       140       150       160       170       180

                  190       200       210       220       230       240
m005.pep   XXXXXXXXXXXXXXASGGYMMACVADKIASAPFAIVGSVGVVAEVPNIHRLLKKHDIDVD
                 :       ||||||||||||:|||||||||||||||||||||||||||||||||
a005       ERNIPLTVAVDKVAASGGYMMACVADKIVSAPFAIVGSVGVVAEVPNIHRLLKKHDIDVD
                  190       200       210       220       230       240

                  250       260       270       280       290       300
m005.pep   VMTAGEFKRTVTFMGENTEKGKQKFRQELEETHQLFKQFVSENRPQLDIEEVATGEHWFG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a005       VMTAGEFKRTVTFMGENTEKGKQKFRQELEETHQLFKQFVSENRPQLDIEEVATGEHWFG
                  250       260       270       280       290       300

                  310       320       330       340       350       360
m005.pep   RQALALNLIDEISTSDDLLLKAFENKQVIEVKYQEKQSLIQRIGLQAEASVEKLFAKLVN
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a005       RQALALNLIDEISTSDDLLLKAFENKQVIEVKYQEKQSLIQRIGLQAEASVEKLFAKLVN
                  310       320       330       340       350       360

m005.pep   RRADVMX
           |||||||
a005       RRADVMX
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 005 shows 77.0% identity over a 366 aa overlap with a predicted ORF (ORF 005.ng) from *N. gonorrhoeae:*

m005/g005

```
                     10        20        30        40        50
m005.pep     MDNIDMFMPEQEEIQSMWKEILLNYGIFLLELLTVFGAIALIVLAIVQSKKQSXSGSV
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
g005       MGMDNIDMFMPEQEEIQSMWKEILLNYGIFLLELLTVFGAIALIVLAIVQSKKQSESGSV
                   10        20        30        40        50        60

           60        70        80        90       100       110
m005.pep   VLTDFSENYKKQRQSFEAFFLSGEEAQHQEKEEKKKEKAEAKAEKKRLKEGGEKSAETXK
           ||||||||||||||||:||||  ||::||||:|||||||||||||||||||||||||| |
g005       VLTDFSENYKKQRQSFETFFLSEEETKHQEKKEKKKEKAEAKAEKKRLKEGGEKSAETQK
                   70        80        90       100       110       120
```

```
             120       130       140       150       160       170
m005.pep      SRLFVLXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
              ||||||                  :
g005          SRLFVLDFDGDLYAHAVESLRHEITAVLLIAKPEDEVLLRLESPGGVVHGYGLAASQLRR
                     130       140       150       160       170       180

             180       190       200       210       220       230
m005.pep      XXXXXXXXXXXXXXXXASGGYMMACVADKIASAPFAIVGSVGVVAEVPNIHRLLKKHDID
                      :       ||||||||||||||:|||||::||||||||||||||||||||||
g005          LRERNIPLTVAVDKVAASGGYMMACVADKIVSAPFAVIGSVGVVAEVPNIHRLLKKHDID
                     190       200       210       220       230       240

             240       250       260       270       280       290
m005.pep      VDVMTAGEFKRTVTFMGENTEKGKQKFRQELEETHQLFKQFVSENRPQLDIEEVATGEHW
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||::||||||
g005          VDVMTAGEFKRTVTFMGENTEKGKQKFRQELEETHQLFKQFVSENRPGLDIEKIATGEHW
                     250       260       270       280       290       300

             300       310       320       330       340       350
m005.pep      FGRQALALNLIDEISTSDDLLLKAFENKQVIEVKYQEKQSLIQRIGLQAEASVEKLFAKL
              ||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
g005          FGRQALALNLIDEISTSDDLLLKAFENKQVIEVKYQEKRSLIQRIGLQAEASVEKLFAKL
                     310       320       330       340       350       360

             360
m005.pep      VNRRADVMX
              |||||||||
g005          VNRRADVMX
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 25>:

g006.seq

```
  1  ATGCTGCTGG TGCTggaatt ttggttCGGc gtGtCGGCGG TGGGCatact
 51  tgCGTTGTTT TTATGGCttt TGCCACGTTT TGCCGCCATC AGCGAAAACC
101  TGTATTTCCG CCTGAACAAC AGCTTGGAAC gcgACAACCA CTTTATCCGA
151  AAAGGCGACG AGCGGCAGCT GTACCGCCAT TACGGACTGG TTTCGCGCCT
201  GCGTGTGCTG ATTTCCAACC GCGAAGCCTT CGGCTATCTC TGCGTCGGCG
251  CGGCGATGGG TATTTTGTTC GGCTTTGCTT TTGTGATGAT GACGCTCAAA
301  GGCTACGGCA GCGCGGGGCA TATTTATTCG GTCGGCACTT ATCTGTGGAT
351  GTTTGCCATG AGTTTGGACG ATGTGCCGCG ATTGGTCGAA CAATATTCCA
401  ATTTGAAAGA CATCGGACAA CGGATAGAGT GGTCGGAACG GAACATCAAA
451  GCCGGAACTT GA
```

This corresponds to the amino acid sequence <SEQ ID 26; ORF 006.ng>:

g006.pep

```
  1  MLLVLEFWFG VSAVGILALF LWLLPRFAAI SENLYFRLNN SLERDNHFIR
 51  KGDERQLYRH YGLVSRLRVL ISNREAFGYL CVGAAMGILF GFAFVMMTLK
101  GYGSAGHIYS VGTYLWMFAM SLDDVPRLVE QYSNLKDIGQ RIEWSERNIK
151  AGT*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 27>:

m006.seq

```
  1  ATGCTGCTGG TGCTGGAATT TTGGGTCGGC GTGTCGGCGG TGGGCATACT
 51  TGCGTTGTTT TTATGGCTTT TGCCACGTTT TGCCGCCATC AGCGAAAACC
101  TGTATTTCCG CCTGAACAAC AGCTTGGAAC GCGACAACCA CTTTATCCGA
151  AAAGGCGACC GGCGGCAGCT GTACCGCCAT TACGGACTGC TTGCGCGCCT
201  GCGTGTGCTG ATTTCCAACC GCGAAGCCTT CGGCTATCTC TGCGTCGGCA
251  CGGCGATGGG TATTTTGTTC GGCTTTGCTT TTGTGATGAT GACGCTCAAA
301  GGCTACAGCA GCGCGGGGCA TGTCTATTCG GTCGGCACTT ATCTGTGGAT
```

```
351  GTTTGCCATG AGTTTGGACG ACGTGCCGCG ATTGGTCGAA CAATATTCCA
401  ATTTGAAAGA CATCGGACAA CGGATAGAGT GGTCGGAACG GAACATCAAA
451  GCCGGAACTTGA
```

This corresponds to the amino acid sequence <SEQ ID 28; ORF 006>:

m006.pep

```
  1  MLLVLEFWVG VSAVGILALF LWLLPRFAAI SENLYFRLNN SLERDNHFIR
 51  KGDRRQLYRH YGLLARLRVL ISNREAFGYL CVGTAMGILF GFAFVMMTLK
101  GYSSAGHVYS VGTYLWMFAM SLDDVPRLVE QYSNLKDIGQ RIEWSERNIK
151  AGT*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 29>:

a006.seq

```
  1  ATGCTGCTGG TGCTGGAATT TTGGGTCGGC GTGTCGGCGG TGGGCATACT
 51  TGCGTTGTTT TTATGGCTTT TGCCACGTTT TGCCGCCATC AGCGAAAACC
101  TGTATTTCCG CCTGAAGAAC AGCTTGGAAC GCGACAACCA CTTTATCCGA
151  AAAGGCGACG AGCGGCAGCT GGACCGCCAT TACGGACTGC TTGCGCGCCT
201  GCGTGTGCTG ATTTCCAACC GCGAAGCCTT CGGCTATCTC TGCGTCGGCA
251  CGGCGATGGG TATTTTGTTC GGCTTTGCTT TTGTGATGAT GACGCTCAAA
301  GGCTACAGCA GCGCGGGGCA TGTCTATTCG GTCGGCACTT ATCTGTGGAT
351  GTTTGCCATA AGTTTGGACG ACGTGCCGCG ATTGGTCGAA CAATATTCCA
401  ATTTGAAAGA CATCGGACAA CGGATAGAGT GGTCGAAACG GAACATCAAA
451  GCCGGAACTT GA
```

This corresponds to the amino acid sequence <SEQ ID 30; ORF 006.a>:

a006.pep

```
  1  MLLVLEFWVG VSAVGILALF LWLLPRFAAI SENLYFRLKN SLERDNHFIR
 51  KGDERQLDRH YGLLARLRVL ISNREAFGYL CVGTAMGILF GFAFVMMTLK
101  GYSSAGHVYS VGTYLWMFAI SLDDVPRLVE QYSNLKDIGQ RIEWSKRNIK
151  AGT*
```

m006/a006 96.7% identity over a 153 aa overlap

```
                     10        20        30        40        50        60
m006.pep    MLLVLEFWVGVSAVGILALFLWLLPRFAAISENLYFRLNNSLERDNHFIRKGDRRQLYRH
            ||||||||||||||||||||||||||||||||||||||||||:|||||||||:|||| ||
a006        MLLVLEFWVGVSAVGILALFLWLLPRFAAISENLYFRLKNSLERDNHFIRKGDERQLDRH
                     10        20        30        40        50        60

                     70        80        90       100       110       120
m006.pep    YGLLARLRVLISNREAFGYLCVGTAMGILFGFAFVMMTLKGYSSAGHVYSVGTYLWMFAM
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
a006        YGLLARLRVLISNREAFGYLCVGTAMGILFGFAFVMMTLKGYSSAGHVYSVGTYLWMFAI
                     70        80        90       100       110       120

                    130       140       150
m006.pep    SLDDVPRLVEQYSNLKDIGQRIEWSERNIKAGTX
            |||||||||||||||||||||||||:||||||||
a006        SLDDVPRLVEQYSNLKDIGQRIEWSKRNIKAGTX
                    130       140       150
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 006 shows 95.4% identity over a 153 aa overlap with a predicted ORF (ORF 006.ng) from *N. gonorrhoeae:*

m006/g006

```
                     10        20        30        40        50        60
m006.pep    MLLVLEFWVGVSAVGILALFLWLLPRFAAISENLYFRLNNSLERDNHFIRKGDRRQLYRH
```

```
            |||||||| |||||||||||||||||||||||||||||||||||||||||||||||:||||||
g006        MLLVLEFWFGVSAVGILALFLWLLPRFAAISENLYFRLNNSLERDNHFIRKGDERQLYRH
                      10        20        30        40        50        60

                      70        80        90       100       110       120
m006.pep    YGLLARLRVLISNREAFGYLCVGTAMGILFGFAFVMMTLKGYSSAGHVYSVGTYLWMFAM
            |||::|||||||||||||||||:||||||||||||||||||||:||||:|||||||||||
g006        YGLVSRLRVLISNREAFGYLCVGAAMGILFGFAFVMMTLKGYGSAGHIYSVGTYLWMFAM
                      70        80        90       100       110       120

                     130       140       150
m006.pep    SLDDVPRLVEQYSNLKDIGQRIEWSERNIKAGTX
            |||||||||||||||||||||||||||||||||
g118        SLDDVPRLVEQYSNLKDIGQRIEWSERNIKAGT
                     130       140       150
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 31>:

```
g006-1.seq
  1   ATGTGGAAAA TGTTGAAACA CATAGCCAAA ACCC[illegible]GCA AGCGATTGAT
 51   TGGCACATTT TCCCCGGTCG GACTGGAAAA CCTT[illegible]GATG CTGGGGTATC
101   CGGTGTTTGG CGGCTGGGCG ATTAATGCCG TGATTGCGGG GAGGGTGTGG
151   CAGGCGTTGC TGTACGCTTT GGTTGTATTT TTGATGTGGC TGGTCGGTGC
201   GGCACGGCGG ATTGCCGATA CGCGCACGTT TACGCGGATT TATACCGAAA
251   TCGCCGTGCC GGTTGTGTTG GAACAACGGC AGCGGCAAGT CCCGCATTCA
301   GCGGTAACTG CACGGGTTGC CCTGTCGCGT GAATTTGTCA GCTTTTTTGA
351   AGAACACCTG CCGATTGCCG CGACATCCGT CGTATCCATA TTCGGCGCGT
401   GCATCATGCT GCTGGTGCTG GAATTTTGGG TCGGCGTGTC GGCGGTGGGC
451   ATACTTGCGT TGTTTTTATG GCTTTTGCCA CGTTTTGCCG CCATCAGCGA
501   AAACCTGTAT TTCCGCCTGA ACAACAGCTT GGAACGCGAC AACCACTTTA
551   TCCGAAAAGG CGACGAGCGG CAGCTGTACC GCCATTACGG ACTGGTTTCG
601   CGCCTGCGTG TGCTGATTTC CAACCGCGAA GCCTTCGGCT ATCTCTGCGT
651   CGGCGCGGCG ATGGGTATTT TGTTCGGCTT TGCTTTTGTG ATGATGACGC
701   TCAAAGGCTA CGGCAGCGCG GGGCATATTT ATTCGGTCGG CACTTATCTG
751   TGGATGTTTG CCATGAGTTT GGACGATGTG CCGCGATTGG TCGAACAATA
801   TTCCAATTTG AAAGACATCG GACAACGGAT AGAGTGGTCG GAACGGAACA
851   TCAAAGCCGG AACTTGA
```

This corresponds to the amino acid sequence <SEQ ID 32; ORF 006-1.ng>:

```
g006-1.pep
  1   MWKMLKHIAK THRKRLIGTF SPVGLENLLM LGYPVFGGWA INAVIAGRVW
 51   QALLYALVVF LMWLVGAARR IADTRTFTRI YTEIAVPVVL EQRQRQVPHS
101   AVTARVALSR EFVSFFEEHL PIAATSVVSI FGACIMLLVL EFWVGVSAVG
151   ILALFLWLLP RFAAISENLY FRLNNSLERD NHFIRKGDER QLYRHYGLVS
201   RLRVLISNRE AFGYLCVGAA MGILFGFAFV MMTLKGYGSA GHIYSVGTYL
251   WMFAMSLDDV PRLVEQYSNL KDIGQRIEWS ERNIKAGT*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 33>:

```
m006-1.seq
  1   ATGTGGAAAA TGTTGAAACA CATAGCCCAA ACCCACCGCA AGCGATTGAT
 51   TGGCACATTT TCCCTGGTCG GACTGGAAAA CCTTTTGATG CTGGTGTATC
101   CGGTGTTTGG CGGCCGGGCG ATCAATGCCG TGATTGCGGG GGAGGTGTGG
151   CAGGCGTTGC TGTACGCTTT GGTTGTGCTT TTGATGTGGC TGGTCGGTGC
201   GGTGCGGCGG ATTGCCGATA CGCGCACGTT TACGCGGATT TATACCGAAA
251   TCGCCGTGCC GGTCGTGTTG GAACAGCGGC AGCGACAAGT CCCGCATTCG
301   GCGGTAACTG CGCGGGTTGC CCTGTCGCGT GAGTTTGTCA GCTTTTTTGA
351   AGAACACCTG CCGATTGCCG CGACATCCGT CGTATCCATA TTCGGCGCGT
401   GCATCATGCT GCTGGTGCTG GAATTTTGGG TCGGCGTGTC GGCGGTGGGC
451   ATACTTGCGT TGTTTTTATG GCTTTTGCCA CGTTTTGCCG CCATCAGCGA
501   AAACCTGTAT TTCCGCCTGA ACAACAGCTT GGAACGCGAC AACCACTTTA
551   TCCGAAAAGG CGACCGGCGG CAGCTGTACC GCCATTACGG ACTGCTTGCG
601   CGCCTGCGTG TGCTGATTTC CAACCGCGAA GCCTTCGGCT ATCTCTGCGT
```

```
651  CGGCACGGCG ATGGGTAT.. TGTTCGGCTT TGCTTTTGTG ATGATGACGC
701  TCAAAGGCTA CAGCAGCGCG GGGCATGTCT ATTCGGTCGG CACTTATCTG
751  TGGATGTTTG CCATGAGTTT GGACGACGTG CCGCGATTGG TCGAACAATA
801  TTCCAATTTG AAAGACATCG GACAACGGAT AGAGTGGTCG GAACGGAACA
851  TCAAAGCCGG AACTTGA
```

This corresponds to the amino acid sequence <SEQ ID 34; ORF 006-1>:

```
m006-1.pep
  1  MWKMLKHIAQ THRKRLIGTF SLVGLENLLM LVYPVFGGRA INAVIAGEVW
 51  QALLYALVVL LMWLVGAVRR IADTRTFTRI YTEIAVPVVL EQRQRQVPHS
101  AVTARVALSR EFVSFFEEHL PIAATSVVSI FGACIMLLVL EFWVGVSAVG
151  ILALFLWLLP RFAAISENLY FRLNNSLERD NHFIRKGDRR QLYRHYGLLA
201  RLRVLISNRE AFGYLCVGTA MGILFGFAFV MMTLKGYSSA GHVYSVGTYL
251  WMFAMSLDDV PRLVEQYSNL KDIGQRIEWS ERNIKAGT*
```

m006-1/g006-1 95.5% identity in 288 aa overlap

```
                    10        20        30        40        50        60
m006-1.pep  MWKMLKHIAQTHRKRLIGTFSLVGLENLLMLVYPVFGGRAINAVIAGEVWQALLYALVVL
            |||||||||:||||||||||| ||||||||| |||||| |||||||||:||||||||||:
g006-1      MWKMLKHIAKTHRKRLIGTFSPVGLENLLMLGYPVFGGWAINAVIAGRVWQALLYALVVF
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m006-1.pep  LMWLVGAVRRIADTRTFTRIYTEIAVPVVLEQRQRQVPHSAVTARVALSREFVSFFEEHL
            |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
g006-1      LMWLVGAARRIADTRTFTRIYTEIAVPVVLEQRQRQVPHSAVTARVALSREFVSFFEEHL
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m006-1.pep  PIAATSVVSIFGACIMLLVLEFWVGVSAVGILALFLWLLPRFAAISENLYFRLNNSLERD
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g006-1      PIAATSVVSIFGACIMLLVLEFWVGVSAVGILALFLWLLPRFAAISENLYFRLNNSLERD
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m006-1.pep  NHFIRKGDRRQLYRHYGLLARLRVLISNREAFGYLCVGTAMGILFGFAFVMMTLKGYSSA
            ||||||||:|||||||||::|||||||||||||||||||:|||||||||||||||||:||
g006-1      NHFIRKGDERQLYRHYGLVSRLRVLISNREAFGYLCVGAAMGILFGFAFVMMTLKGYGSA
                   190       200       210       220       230       240

                   250       260       270       280      289
m006-1.pep  GHVYSVGTYLWMFAMSLDDVPRLVEQYSNLKDIGQRIEWSERNIKAGTX
            ||:||||||||||||||||||||||||||||||||||||||||||||||
g006-1      GHIYSVGTYLWMFAMSLDDVPRLVEQYSNLKDIGQRIEWSERNIKAGTX
                   250       260       270       280
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 35>:

```
a006-1.seq  (partial)
  1  ..AGCCAAAACC ACCGCAAGCG ATTGATTGGC ACATTTTTTC TGGTCGGACT
 51    GGAAAACCTT TTGATGCTGG TGTATCCGGT GTTTGGCGGC TGGGCGATTA
101    ATGCCGTGAT TGCGGGGCAG GCGTGGCAGG CGTTGCTGTA CGCTTTGGTT
151    GTGCTTTTGA TGTGGCTGGT CGGTGCGGCG CGGCGGATTG CCGATACGCG
201    CACGTTTACG CGGATTTATA CCGAAATCGC CGTGCCGGTT GTGTTGGAAC
251    AGCGGCAGCG GCAAGTCCCG CATTCGGCGG TAACTGCGCG GGTTGCCCTG
301    TCGCGTGAGT TTGTCAGCTT TTTTGAAGAA CACCTGCCGA TTGCCGCGAC
351    ATCCGTCGTA TCCATATTCG GCGCGTGCAT CATGCTGCTG GTGCTGGAAT
401    TTTGGGTCGG CGTGTCGGCG GTGGGCATAC TTGCGTTGTT TTTATGGCTT
451    TTGCCACGTT TTGCCGCCAT CAGCGAAAAC CTGTATTTCC GCCTGAAGAA
501    CAGCTTGGAA CGCGACAACC ACTTTATCCG AAAAGGCGAC GAGCGGCAGC
551    TGGACCGCCA TTACGGACTG CTTGCGCGCC TGCGTGTGCT GATTTCCAAC
601    CGCGAAGCCT TCGGCTATCT CTGCGTCGGC ACGGCGATGG GTATTTTGTT
651    CGGCTTTGCT TTTGTGATGA TGACGCTCAA AGGCTACAGC AGCGCGGGGC
```

```
701   ATGTCTATTC GGTCGGCACT TATCTGTGGA TGTTTGCCAT AAGTTTGGAC
751   GACGTGCCGC GATTGGTCGA ACAATATTCC AATTTGAAAG ACATCGGACA
801   ACGGATAGAG TGGTCGAAAC GGAACATCAA AGCCGGAACT TGA
```

This corresponds to the amino acid sequence <SEQ ID 36; ORF 006-1.a>:

```
a006-1.pep (partial)
  1   ..SQNHRKRLIG TFFLVGLENL LMLVYPVFGG WAINAVIAGQ AWQALLYALV
 51   VLLMWLVGAA RRIADTRTFT RIYTEIAVPV VLEQRQRQVP HSAVTARVAL
101   SREFVSFFEE HLPIAATSVV SIFGACIMLL VLEFWVGVSA VGILALFLWL
151   LPRFAAISEN LYFRLKNSLE RDNHFIRKGD ERQLDRHYGL LARLRVLISN
201   REAFGYLCVG TAMGILFGFA FVMMTLKGYS SAGHVYSVGT YLWMFAISLD
251   DVPRLVEQYS NLKDIGQRIE WSKRNIKAGT *
```

```
a006-1/m006-1   95.7% identity in 280 aa overlap
                          10        20        30        40        50
a006-1.pep            SQNHRKRLIGTFFLVGLENLLMLVYPVFGGWAINAVIAGQAWQALLYALVVL
                      :|:||||||||| ||||||||||||||||||| ||||||||::||||||||||||
m006-1          MWKMLKHIAQTHRKRLIGTFSLVGLENLLMLVYPVFGGRAINAVIAGEVWQALLYALVVL
                        10        20        30        40        50        60

                  60        70        80        90       100       110
a006-1.pep      LMWLVGAARRIADTRTFTRIYTEIAVPVVLEQRQRQVPHSAVTARVALSREFVSFFEEHL
                |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
m006-1          LMWLVGAVRRIADTRTFTRIYTEIAVPVVLEQRQRQVPHSAVTARVALSREFVSFFEEHL
                        70        80        90       100       110       120

                 120       130       140       150       160       170
a006-1.pep      PIAATSVVSIFGACIMLLVLEFWVGVSAVGILALFLWLLPRFAAISENLYFRLKNSLERD
                |||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
m006-1          PIAATSVVSIFGACIMLLVLEFWVGVSAVGILALFLWLLPRFAAISENLYFRLNNSLERD
                       130       140       150       160       170       180

                 180       190       200       210       220       230
a006-1.pep      NHFIRKGDERQLDRHYGLLARLRVLISNREAFGYLCVGTAMGILFGFAFVMMTLKGYSSA
                ||||||||:||| |||||||||||||||||||||||||||||||||||||||||||||||
m006-1          NHFIRKGDRRQLYRHYGLLARLRVLISNREAFGYLCVGTAMGILFGFAFVMMTLKGYSSA
                       190       200       210       220       230       240

                 240       250       260       270       280
a006-1.pep      GHVYSVGTYLWMFAISLDDVPRLVEQYSNLKDIGQRIEWSKRNIKAGTX
                ||||||||||||||:|||||||||||||||||||||||||:||||||||
m006-1          GHVYSVGTYLWMFAMSLDDVPRLVEQYSNLKDIGQRIEWSERNIKAGTX
                       250       260       270       280
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 37>:

```
g007.seq
  1   atgaACACAA CCCGACTGCC GACCGCCTTC ATCTTGTGCT GCCTCTGcgC
 51   CGCcGCTTCT GCCGccgaca acAGCatcat gaCaAAAGGG CAAAAAGTGT
101   ACGAATCcAa ctGCATCGCC TGCCACGGCA AGAAAGGGGA AGGGCGCGGC
151   ACTGCGtTTC CTccgctTTT CCggtcgGac tgtattatga acaAACCGCa
201   cgTCCtgctg cacagcatgg tcaaaggcAt cgacgggaca ttcaaagtgg
251   agcggcaaaa cctacgacgg atttatgCcc gcaaccgcca tcagcgATGC
301   GGACATTGCC GCCGTCGCCA CTTATATCAT GAACGCCTTT GA
```

This corresponds to the amino acid sequence <SEQ ID 38; ORF 007.ng>:

```
g007.pep
  1   MNTTRLPTAF ILCCLCAAAS AADNSIMTKG QKVYESNCIA CHGKKGEGRG
 51   TAFPPLFRSD CIMNKPHVLL HSMVKGIDGT FKVERQNLRR IYARNRHQRC
101   GHCRRRHLYH ERL*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 39>:

m007.seq

```
  1  ATGAACACAA CCCGACTGCC GACCGCCCTC GTCTTGGGCT GCTTCTGCGC
 51  CGCCGCTTCT GCCGCCGACA ACAGCATCAT GACAAAAGGG CAAAAAGTGT
101  ACGAATCCAA CTGCGTCGCC TGCCACGGCA AAAAGGGCGA AGGCCGCGGA
151  ACCATGTTTC CGCCGCTCTA CCGCTCCGAC TTCATCATGA AAAAACCGCA
201  GGTGCTGCTG CACAGCATGG TCAAAGGCAT CAACGGTACA ATCAAAGTC.
251  AACGGCAAAA CCTACAACGG ATTCATGCCC GCAACCGCCA TCAGCGATGC
301  GGACATTGCC GCCGTCGCCA CTTATATCAT GAACGCCTTT GA
```

This corresponds to the amino acid sequence <SEQ ID 40; ORF 007>:

m007.pep

```
  1  MNTTRLPTAL VLGCFCAAAS AADNSIMTKG QKVYESNCVA CHGKKGEGRG
 51  TMFPPLYRSD FIMKKPQVLL HSMVKGINGT IKVXRQNLQR IHARNRHQRC
101  GHCRRRHLYH ERL*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 41>:

a007.seq

```
  1  ATGAACACAA CCCGACTGCC GACCGCCCTC GTCTTGGGCT GCCTCTGCGC
 51  CGCCGCTTCT GCCGCCGACA ACAGCATCAT GACAAAAGGG CAAAAAGTGT
101  ACGAATCCAA CTGCGTCGCC TGCCACGGCA AAAAGGGCGA AGGCCGCGGA
151  ACCATGTTTC CGCCGCTCTA CCGCTCCGAC TTCATCATGA AAAAACCGCA
201  GGTGCTGCTG CACAGCATGG TCAAAGGCAT CAACGGTACA ATCAAAGTC.
251  AACGGCAAAA CCTACAACGG ATTCATGCCC GCCACTGCCA TCAGCGATGC
301  GGACATTGCC GCCGTCGCCA CTTATATCAT GAACGCCTTT GA
```

This corresponds to the amino acid sequence <SEQ ID 42; ORF 007.a>:

a007.pep

```
  1  MNTTRLPTAL VLGCLCAAAS AADNSIMTKG QKVYESNCVA CHGKKGEGRG
 51  TMFPPLYRSD FIMKKPQVLL HSMVKGINGT IKVXRQNLQR IHARHCHQRC
101  GHCRRRHLYH ERL*
```

m007/a007 97.3% identity over a 113 aa overlap

```
                      10        20        30        40        50        60
m007.pep      MNTTRLPTALVLGCFCAAASAADNSIMTKGQKVYESNCVACHGKKGEGRGTMFPPLYRSD
              |||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
a007          MNTTRLPTALVLGCLCAAASAADNSIMTKGQKVYESNCVACHGKKGEGRGTMFPPLYRSD
                      10        20        30        40        50        60

                      70        80        90       100       110
m007.pep      FIMKKPQVLLHSMVKGINGTIKVXRQNLQRIHARNRHQRCGHCRRRHLYHERLX
              ||||||||||||||||||||||||||||||||||: ||||||||||||||||||
a007          FIMKKPQVLLHSMVKGINGTIKVXRQNLQRIHARHCHQRCGHCRRRHLYHERLX
                      70        80        90       100       110
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

ORF 007 shows 86.7% identity over a 113 aa overlap with a predicted ORF (ORF 007.ng) from *N. gonorrhoeae:*

m007/g007

```
                      10        20        30        40        50        60
m007.pep      MNTTRLPTALVLGCFCAAASAADNSIMTKGQKVYESNCVACHGKKGEGRGTMFPPLYRSD
              |||||||||::| |:|||||||||||||||||||||||:|||||||||||| ||||:|||
g007          MNTTRLPTAFILCCLCAAASAADNSIMTKGQKVYESNCIACHGKKGEGRGTAFPPLFRSD
                      10        20        30        40        50        60

                      70        80        90       100       110
m007.pep      FIMKKPQVLLHSMVKGINGTIKVXRQNLQRIHARNRHQRCGHCRRRHLYHERLX
              ||:||:||||||||||:||:|| ||||:||:||||||||||||||||||||||
```

```
g007          CIMNKPHVLLHSMVKGIDGTFKVERQNLRRIYARNRHQRCGHCRRRHLYHERL
                        70        80        90       100       110
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 43>:

```
g007-1.seq  (partial)
     1  ATGAACACAA CCCGACTGCC GACCGCCTTC ATCTTGTGCT GCCTCTGCGC
    51  CGCCGCTTCT GCCGCCGACA ACAGCATCAT GACAAAAGGG CAAAAAGTGT
   101  ACGAATCCAA CTGCATCGCC TGCCACGGCA AGAAAGGGGA AGGGCGCGGC
   151  ACTGCGTTTC CTCCGCTTTT CCGGTCGGAC TATATTATGA ACAAACCGCA
   201  CGTCCTGCTG CACAGCATGG TCAAAGGCAT CAACGGTACA ATCAAAGTCA
   251  ACGGCAAAAC CTACAACGGA TTCATGCCCG CAACCGCCAT CAGCGATGCG
   301  GACATTGCCG CCGTCGCCAC TTATATCATG AACGCCTTTG ACAACGGCGG
   351  CGGAAGCGTT ACCGAAAAAG ACGTAAAACA GGCAAAAGGC AAAAAAAAC.
```

This corresponds to the amino acid sequence <SEQ ID 44; ORF 007-1.ng>:

```
g007-1.pep  (partial)
     1  MNTTRLPTAF ILCCLCAAAS AADNSIMTKG QKVYESNCIA CHGKKGEGRG
    51  TAFPPLFRSD YIMNKPHVLL HSMVKGINGT IKVNGKTYNG FMPATAISDA
   101  DIAAVATYIM NAFDNGGGSV TEKDVKQAKG KKN...
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 45>:

```
m007-1.seq
     1  ATGAACACAA CCCGACTGCC GACCGCCCTC GTCTTGGGCT GCTTCTGCGC
    51  CGCCGCTTCT GCCGCCGACA ACAGCATCAT GACAAAAGGG CAAAAAGTGT
   101  ACGAATCCAA CTGCGTCGCC TGCCACGGCA AAAAGGGCGA AGGCCGCGGA
   151  ACCATGTTTC CGCCGCTCTA CCGCTCCGAC TTCATCATGA AAAAACCGCA
   201  GGTGCTGCTG CACAGCATGG TCAAAGGCAT CAACGGTACA ATCAAAGTCA
   251  ACGGCAAAAC CTACAACGGA TTCATGCCCG CAACCGCCAT CAGCGATGCG
   301  GACATTGCCG CCGTCGCCAC TTATATCATG AACGCCTTTG ACAACGGCGG
   351  CGGAAGCGTT ACCGAAAAAG ACGTAAAACA GGCAAAAAGC AAAAAAAACT
   401  AA
```

This corresponds to the amino acid sequence <SEQ ID 46; ORF 007-1>

```
m007-1.pep
     1  MNTTRLPTAL VLGCFCAAAS AADNSIMTKG QKVYESNCVA CHGKKGEGRG
    51  TMFPPLYRSD FIMKKPQVLL HSMVKGINGT IKVNGKTYNG FMPATAISDA
   101  DIAAVATYIM NAFDNGGGSV TEKDVKQAKS KKN*
```

m007-1 / g007-1   91.7% identity in 133 aa overlap

```
                    10        20        30        40        50        60
m007-1.pep  MNTTRLPTALVLGCFCAAASAADNSIMTKGQKVYESNCVACHGKKGEGRGTMFPPLYRSD
            |||||||||::| |:||||||||||||||||||||||||:||||||||||| ||||:|||
g007-1      MNTTRLPTAFILCCLCAAASAADNSIMTKGQKVYESNCIACHGKKGEGRGTAFPPLFRSD
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m007-1.pep  FIMKKPQVLLHSMVKGINGTIKVNGKTYNGFMPATAISDADIAAVATYIMNAFDNGGGSV
            :||:||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
g007-1      YIMNKPHVLLHSMVKGINGTIKVNGKTYNGFMPATAISDADIAAVATYIMNAFDNGGGSV
                    70        80        90       100       110       120

                   130
m007-1.pep  TEKDVKQAKSKKNX
            |||||||||:|||
g007-1      TEKDVKQAKGKKN
                   130
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 47>:

```
a007-1.seq  (partial)
     1  ATGAACACAA CCCGACTGCC GACCGCCCTC GTCTTGGGCT GCCTCTGCGC
    51  CGCCGCTTCT GCCGCCGACA ACAGCATCAT GACAAAAGGG CAAAAAGTGT
   101  ACGAATCCAA CTGCGTCGCC TGCCACGGCA AAAAGGGCGA AGGCCGCGGA
```

```
151  ACCATGTTTC CGCCGCTC.A CCGCTCCGAC TTCATCATGA AAAAACCGCA
201  GGTGCTGCTG CACAGCATGG TCAAAGGCAT CAACGGTACA ATCAAAGTCA
251  ACGGCAAAAC CTACAACGGA TTCATGCCCG CCACTGCCAT CAGCGATGCG
301  GACATTGCCG CCGTCGCCAC TTATATCATG AACGCCTTTG ACAACGGCGG
351  CGGAAGCGTT ACCGAAAAAG ACGTAAAACA GGCAAAAAAC AAAAAA..
```

This corresponds to the amino acid sequence <SEQ ID 48; ORF 007-1.a>:

a007-1.pep (partial)

```
  1  MNTTRLPTAL VLGCLCAAAS AADNSIMTKG QKVYESNCVA CHGKKGEGRG
 51  TMFPPLYRSD FIMKKPQVLL HSMVKGINGT IKVNGKTYNG FMPATAISDA
101  DIAAVATYIM NAFDNGGGSV TEKDVKQAKN KK..
```

m007-1/a007-1 98.5% identity in 132 aa overlap

```
                        10        20        30        40        50        60
m007-1.pep      MNTTRLPTALVLGCFCAAASAADNSIMTKGQKVYESNCVACHGKKGEGRGTMFPPLYRSD
                ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
a007-1          MNTTRLPTALVLGCLCAAASAADNSIMTKGQKVYESNCVACHGKKGEGRGTMFPPLYRSD
                        10        20        30        40        50        60

                        70        80        90       100       110       120
m007-1.pep      FIMKKPQVLLHSMVKGINGTIKVNGKTYNGFMPATAISDADIAAVATYIMNAFDNGGGSV
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a007-1          FIMKKPQVLLHSMVKGINGTIKVNGKTYNGFMPATAISDADIAAVATYIMNAFDNGGGSV
                        70        80        90       100       110       120

                       130
m007-1.pep      TEKDVKQAKSKKNX
                |||||||||:||
a007-1          TEKDVKQAKNKK
                       130
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 49>:

g008.seq

```
  1  ATGAACAACA GACATTTTGC CGTCAtcgCC TTGGGCAGCA ACCTTGACAA
 51  CCCCGCACAA CAAATacgcg gcgcattaga cgcgctctcg tcccatcctg
101  acatccggct tgaaCaggtt tcctcactgt aTatgaccgc acctgtcggt
151  tacgAcaaTC agcccgATTT CATCaatgcc gTCTgcaccg TTTCCACCAC
201  CtTGGACGGC ATTGcccTGC TTGCCgaACT CAAccgTATC GAAGCCGATT
251  TCGGACGCGA aCGCAGTTTC CGCAATGCAC CGCGCACATT GGATTTGGAC
301  ATTATCGACT TTGACGGCAT CTCCAGCGAC GACCCCCGCC TTACCCTGCC
351  GCATCCGCGC GCGCACGAAC GCAGTTTCGT CATACGCCCT TTGGCAGAAA
401  TCCTCCCTGA TTTTATTTTG GGAAAATACG GAAAGGTTGT CGAATTGTCA
451  AAACGGCTGG GCAATCAAGG CATCCGTCTT TTACCGGACA GGTAA
```

This corresponds to the amino acid sequence <SEQ ID 50; ORF 008.ng>:

g008.pep

```
  1  MNNRHFAVIA LGSNLDNPAQ QIRGALDALS SHPDIRLEQV SSLYMTAPVG
 51  YDNQPDFINA VCTVSTTLDG IALLAELNRI EADFGRERSF RNAPRTLDLD
101  IIDFDGISSD DPRLTLPHPR AHERSFVIRP LAEILPDFIL GKYGKVVELS
151  KRLGNQGIRL LPDR*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 51>:

m008.seq

```
  1  ATGAACAACA GACATTTTGC CGTCATCGCC CTGGGCAGTA ATCTTGAAAA
 51  CCCTGCTCAA CAGGTACGCG CCGCATTGGA CACGCTGTCG TCCCATCCTG
101  ACATCCGTCT TAAACAGGCT TCCTCACTGT ATATGACCGC GCCCGTCGGT
151  TACGACAATC AGCCCGATTT TGTCAATGCC GTCTGCACCG TTTCCACCAC
201  TCTGGACGGC ATTGCCyTGC TTGCCGAACT CAACCGTATC GAGGCTGATT
251  TCGGACGCGA ACGCAGCTTC CGCAACGCGC CGCGCACATT GkATTTGGAC
301  ATTATCGACT TTGACGGCAT CTCCAGCGAC GACACsCGAC TcACCtTGCC
```

```
351  GCATCCGCGC GCGCACGAAC GCAGTTTCGT CATCCGCCCT TTGGCAGAAA
401  TCCTCCCTGA TTTTGTTTTA GGAAAACACG GAAAGGTTGC CGAATTGTCA
451  AAACGGyTGG GCAATCAAGG TATCCGTCTT TTACCGGACA GGTAATT
```

This corresponds to the amino acid sequence <SEQ ID 52; ORF 008>:

m008.pep

```
  1  MNNRHFAVIA LGSNLENPAQ QVRAALDTLS SHPDIRLKQA SSLYMTAPVG
 51  YDNQPDFVNA VCTVSTTLDG IALLAELNRI EADFGRERSF RNAPRTLXLD
101  IIDFDGISSD DTRLTLPHPR AHERSFVIRP LAEILPDFVL GKHGKVAELS
151  KRLGNQGIRL LPDR*
```

The following partial DNA sequence was identified in *N. meningitidis*<SEQ ID 53>:

a008.seq

```
  1  ATGAACAACA GACATTTTGC CGTCATCGCC CTGGGCAGTA ATCTTGAAAA
 51  CCCTGCCCAA CAGGTACGCG CCGCATTGGA CACGCTGTCG TCCCATCCTG
101  ACATCCGTCT TAAACAGGCT TCCTCACTGT ATATGACCGC GCCCGTCGGT
151  TACGACAATC AGCCCGATTT CGTCAATGCC GTCTGCACCG TTTCCACCAC
201  CTTGGACGGC ATTGCCCTGC TTGCCGAACT CAACCGTATC GAAGCCGATT
251  TCGGACGCGA ACGCAGCTTC CGCAACGCGC CGCGCACATT GGATTTGGAC
301  ATTATCGACT TTGACGGCAT CTCCAGCGAC GACCCCCGAC TCACCCTGCC
351  GCATCCGCGC GCGCACGAAC GCAGTTTCGT CATACGCCCT TTGGCAGAAA
401  TCCTCCCTGA TTTTATTTTG GGAAAACACG GAAAGGTTGC CGAATTGTCA
451  AAACGGCTGG GCAATCAAGG CATCCGTCTT TTACCGGATA AGTAA
```

This corresponds to the amino acid sequence <SEQ ID 54; ORF 008.a>:

a008.pep

```
  1  MNNRHFAVIA LGSNLENPAQ QVRAALDTLS SHPDIRLKQA SSLYMTAPVG
 51  YDNQPDFVNA VCTVSTTLDG IALLAELNRI EADFGRERSF RNAPRTLDLD
101  IIDFDGISSD DPRLTLPHPR AHERSFVIRP LAEILPDFIL GKHGKVAELS
151  KRLGNQGIRL LPDK*
```

m008/a008 97.6% identity over a 164 aa overlap

```
                     10        20        30        40        50        60
m008.pep     MNNRHFAVIALGSNLENPAQQVRAALDTLSSHPDIRLKQASSLYMTAPVGYDNQPDFVNA
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a008         MNNRHFAVIALGSNLENPAQQVRAALDTLSSHPDIRLKQASSLYMTAPVGYDNQPDFVNA
                     10        20        30        40        50        60

                     70        80        90       100       110       120
m008.pep     VCTVSTTLDGIALLAELNRIEADFGRERSFRNAPRTLXLDIIDFDGISSDDTRLTLPHPR
             ||||||||||||||||||||||||||||||||||||||||| ||||||||| ||||||||
a008         VCTVSTTLDGIALLAELNRIEADFGRERSFRNAPRTLDLDIIDFDGISSDDPRLTLPHPR
                     70        80        90       100       110       120

                    130       140       150       160
m008.pep     AHERSFVIRPLAEILPDFVLGKHGKVAELSKRLGNQGIRLLPDRX
             ||||||||||||||||||:|||||||||||||||||||||||:|
a008         AHERSFVIRPLAEILPDFILGKHGKVAELSKRLGNQGIRLLPDKX
                    130       140       150       160
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

ORF 008 shows 92.7% identity over a 164 aa overlap with a predicted ORF (ORF008.ng) from *N. gonorrhoeae:*

m008/g008

```
                     10        20        30        40        50        60
m008.pep     MNNRHFAVIALGSNLENPAQQVRAALDTLSSHPDIRLKQASSLYMTAPVGYDNQPDFVNA
             |||||||||||||||:|||||:|:|||:|||||||||:|:|||||||||||||||||:||
g008         MNNRHFAVIALGSNLDNPAQQIRGALDALSSHPDIRLEQVSSLYMTAPVGYDNQPDFINA
```

```
                       10        20        30        40        50        60

                       70        80        90       100       110       120
m008.pep     VCTVSTTLDGIALLAELNRIEADFGRERSFRNAPRTLXLDIIDFDGISSDDTRLTLPHPR
             ||||||||||||||||||||||||||||||||||||| ||||||||||||| ||||||||
g008         VCTVSTTLDGIALLAELNRIEADFGRERSFRNAPRTLDLDIIDFDGISSDDPRLTLPHPR
                       70        80        90       100       110       120

                      130       140       150       160
m008.pep     AHERSFVIRPLAEILPDFVLGKHGKVAELSKRLGNQGIRLLPDRX
             ||||||||||||||||||:|||:|||:||||||||||||||||||
g008         AHERSFVIRPLAEILPDFILGKYGKVVELSKRLGNQGIRLLPDRX
                      130       140       150       160
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 55>:

g009.seq

```
  1  ATGCCCCGCG CTGCCGTAGC CTTTGAGCGT CATCATCACA AAAGCAAAGC
 51  CGAACAAAAT ACCCATCGCC GCGCCGACGC AGAGATAGCC GAAGGCTTCG
101  CGGTTGGAAA TCAGCACACG CAGGCGCGAA ACCAGTCCGT AATGGCGGTA
151  CAGCTGCCGC TCGTCGCCTT TTCGGATAAA GTGGTTGTcg cGTTCCAAGC
201  TGTTGTTCAG GCGGAAATAC AGGTTTTCGC TGATGGCGGC AAAACGTGGC
251  AaaaGCCATA A
```

This corresponds to the amino acid sequence <SEQ ID 56; ORF 009.ng>:

g009.pep

```
  1  MPRAAVAFER HHHKSKAEQN THRRADAEIA EGFAVGNQHT QARNQSVMAV
 51  QLPLVAFSDK VVVAFQAVVQ AEIQVFADGG KTWQKP*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 57>:

m009.seq

```
  1  ATGCCCCGCG CTGCTGTAGC CTTTGAGCGT CATCATCACA AAAGCAAAGC
 51  CGAACAAAAT ACCCATCGCC GTGCCGACGC AGAGATAGCC GAAGGCTTCG
101  CGGTTGGAAA TCAGCACACG CAGGCGCGCA AGCAGTCCGT AATGGCGGTA
151  CAGCTGCCGC CGGTCGCCTT TTCGGATAAA GTGGTTGTCG CGTTCCAAGC
201  TGTTGTTCAG GCGGAAATAC AGGTTTTCGC TGATGGCGGC AAAACGTGGC
251  AAAAGCCATA A
```

This corresponds to the amino acid sequence <SEQ ID 58; ORF 009>:

m009.pep

```
  1  MPRAAVAFER HHHKSKAEQN THRRADAEIA EGFAVGNQHT QARKQSVMAV
 51  QLPPVAFSDK VVVAFQAVVQ AEIQVFADGG KTWQKP*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 009 shows 97.7% identity over a 86 aa overlap with a predicted ORF (ORF 009.ng) from *N. gonorrhoeae:*

m009/g009

```
                       10        20        30        40        50        60
m009.pep     MPRAAVAFERHHHKSKAEQNTHRRADAEIAEGFAVGNQHTQARKQSVMAVQLPPVAFSDK
             |||||||||||||||||||||||||||||||||||||||||||:||||||||| ||||||
g009         MPRAAVAFERHHHKSKAEQNTHRRADAEIAEGFAVGNQHTQARNQSVMAVQLPLVAFSDK
                       10        20        30        40        50        60

                       70        80
m009.pep     VVVAFQAVVQAEIQVFADGGKTWQKPX
             |||||||||||||||||||||||||||
g009         VVVAFQAVVQAEIQVFADGGKTWQKPX
                       70        80
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 59>:

```
a009.seq
     1  ATGCCCCGCG CTGCTGTAGC CTTTGAGCGT CATCATCACA AAAGCAAAGC
    51  CGAACAAAAT ACCCATCGCC GTGCCGACGC AGAGATAGCC GAAGGCTTCG
   101  CGGTTGGAAA TCAGCACACG CAGGCGCGCA AGCAGTCCGT AATGGCGGTC
   151  CAGCTGCCGC TCGTCGCCTT TTCGGATAAA GTGGTTGTCG CGTTCCAAGC
   201  TGTTCTTCAG GCGGAAATAC AGGTTTTCGC TGATGGCGGC AAAACGTGGC
   251  AAAAGCCATA A
```

This corresponds to the amino acid sequence <SEQ ID 60; ORF 009.a>:

```
a009.pep
     1  MPRAAVAFER HHHKSKAEQN THRRADAEIA EGFAVGNQHT QARKQSVMAV
    51  QLPLVAFSDK VVVAFQAVLQ AEIQVFADGG KTWQKP*
```

m009/a009 97.7% identity over a 86 aa overlap

```
                        10        20        30        40        50        60
m009.pep        MPRAAVAFERHHHKSKAEQNTHRRADAEIAEG[illegible]GNQHTQARKQSVMAVQLPPVAFSDK
                ||||||||||||||||||||||||||||||||||.||||||||||||||||||| |||||
a009            MPRAAVAFERHHHKSKAEQNTHRRADAEIAEGFAVGNQHTQARKQSVMAVQLPLVAFSDK
                        10        20        30        40        50        60

                        70        80
m009.pep        VVVAFQAVVQAEIQVFADGGKTWQKPX
                ||||||||:||||||||||||||||||
a009            VVVAFQAVLQAEIQVFADGGKTWQKPX
                        70        80
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 61>:

```
g010.seq
     1  ATGGGTTTTC CTGTTCGCAA GTTTGATGCC GTGATTGTCG GCGGTGGCGG
    51  TGCAGGTTTA CGTGCAGCCC TCCAATTATC CAAATCCGGT TTGAATTGTG
   101  CCGTTTTGTC TAAAGTGTTC CCGACCCGCT CGCATACCGT AGCGGCGCAG
   151  GGCGGTATTT CCGCCTCTCT GGGTAATGTG CAGGAGGACC GTTGGGACTG
   201  GCACATGTAC GATACCGTGA AAGGTTCCGA CTGGCTGGGC GACCAAGATG
   251  CGATTGAGTT TATGTGTCGC GCTGCGCCTG AAGCGGTGAT TGAGTTGGAA
   301  CACATGGGTA TGCCTTTTGA CCGCGTTGAA AGCGGCAAAA TTTATCAGCG
   351  TCCTTTCGGC GGACATACTG CCGAACATGG TAAACGTGCG GTAGAACGTG
   401  CATGTGCGGT TGCCGACCGT ACCGGTCATG CGATGTTGCA TACTTTGTAC
   451  CAACAAAACG TCCGTGCCAA TACACAATTC TTTGTGGAAT GGACGGCGCA
   501  AGATTTGATT CGTGATGAAA ACGGCGATGT CGTCGGCGTA ACCGCCATGG
   551  AAATGGAAAC GGGCGAAGTT TATATTTTCC ACGCCAAGGC CGTGATGTTT
   601  GCTACCGGTG GCGGCGGTCG TATTTATGCT TCTTCTACCA ATGCTTATAT
   651  GAATACCGGT GACGGTTTGG GCATTTGCGC CCGTGCGGGC ATTCCGTTGG
   701  AAGATATGGA ATTCTGGCAA TTCCACCCGA CCGGCGTGGC GGGTGCGGGC
   751  GTGTTGATTA CCGAAGGCGT ACGCGGCGAG GGCGGTATTC TGTTGAacgc
   801  cgacggcgaA cgcTTTATGG AAcgctatgc GCcgACCGta aAagaCTTGG
   851  CTTCTCGCga cgtGGTTTCA CgcgcGatgG CGatggaAAt ctatgaaggt
   901  cgcggctgTG GtaaAAAcaA agaCCacgtC TTACTGAAAA TCGACcAtAt
   951  cggtGCAGAA AAAATTATGG AAAAACTGCC GGGCATCCGC GAGATTTCCA
  1001  TTCagtttgc cGGTATCGAT CCGATTAAAG ACCCGATTcc ggttgTGCCG
  1051  ACTACCCACT ATATGATGGG CGGCATTCcg aCCAATTATC ACGGTGAAGT
  1101  TGTTGTTCCG CAAGGCGACG AGTACGAAGT ACCTGTAAAA GGCCTGTATG
  1151  CCGCAGGTGA GTGCGCCTGT GCTTCCGTAC ACGGTGCGAA CCGTTTGGGT
  1201  ACGAACTCCC TGCTGGACTT GGTGGTGTTC cgcccaaccc cccggtga
```

This corresponds to the amino acid sequence <SEQ ID 62; ORF 010.ng>:

g010.pep

```
  1  MGFPVRKFDA VIVGGGGAGL RAALQLSKSG LNCAVLSKVF PTRSHTVAAQ
 51  GGISASLGNV QEDRWDWHMY DTVKGSDWLG DQDAIEFMCR AAPEAVIELE
101  HMGMPFDRVE SGKIYQRPFG GHTAEHGKRA VERACAVADR TGHAMLHTLY
151  QQNVRANTQF FVEWTAQDLI RDENGDVVGV TAMEMETGEV YIFHAKAVMF
201  ATGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG
251  VLITEGVRGE GGILLNADGE RFMERYAPTV KDLASRDVVS RAMAMEIYEG
301  RGCGKNKDHV LLKIDHIGAE KIMEKLPGIR EISIQFAGID PIKDPIPVVP
351  TTHYMMGGIP TNYHGEVVVP QGDEYEVPVK GLYAAGECAC ASVHGANRLG
401  TNSLLDLVVF RPTPR*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 63>:

m010.seq (PARTIAL)

```
  1  ..nTCCAATTAT CCAAATCCGG TCTGAATTGT GCCGTTTTGT CTAAAGTGTT
 51    CCCGACCCGT TCGCATACCG TAGCGGCGCA GGGCGGTATT TCCGCCTCTn
101    TGGGTAATGT GCAGGAAGAC CGTTGGGACT GGCACATGTA CGATACCGTG
151    AAAGGTTCCG ACTGGTTGGG CGACCAAGAT GCGATTGAGT TTATGTGCCG
201    CGCCGCGCCT GAAGCCGTAA TTGAGTTGGA ACACATGGGT ATGCCTTTTG
251    ACCGTGTGGA AAGCGGTAAA ATTTATCAGC GTCCTTTCGG CGGCCATACT
301    GCCGAACACG GTAAACGCGC GGTAGAACGC GyCTGTGCGG TTGCCGACCG
351    TACAGGTCAT GCGATGCTGC ATACTTTGTA CCAACAAAAC GTCCGTGCCA
401    ATACGCAATT CTTTGTGGAA TGGACGGCAC AAGATTTGAT TCGTGATGAA
451    AACGGCGATG TCGTCGGCGT AACCGCCATG GAAATGGAAA CCGGCGAAgT
501    TTATATTTTC CACGCTAAAG CTGTGATGTT TGCTACCGGC GGCGGCGGTC
551    GTATTTATGC GTCTTCTACC AATGCCTATA TGAATACCGG CGATGGTTTG
601    GGTATTTGTG CGCGTGCAGG TATCCCGTTG GAAGACATGG AATTCTGGCA
651    ATTCCAGCCG ACCGGCGTGG CGGGTGCGGG CGTGTTGATT ACCGAA....
```

This corresponds to the amino acid sequence <SEQ ID 64; ORF 010>:

m010.pep (PARTIAL)

```
  1  ..XQLSKSGLNC AVLSKVFPTR SHTVAAQGGI SASXGNVQED RWDWHMYDTV
 51    KGSDWLGDQD AIEFMCRAAP EAVIELEHMG MPFDRVESGK IYQRPFGGHT
101    AEHGKRAVER XCAVADRTGH AMLHTLYQQN VRANTQFFVE WTAQDLIRDE
151    NGDVVGVTAM EMETGEVYIF HAKAVMFATG GGGRIYASST NAYMNTGDGL
201    GICARAGIPL EDMEFWQFQP TGVAGAGVLI TE...
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 65>:

a010.seq

```
   1  ATGGGCTTTC CTGTTCGCAA GTTTGATGCC GTGATTGTCG GCGGTGGTGG
  51  TGCAGGTTTA CGCGCANCCC TCCAATTATC CAAATCCGGT CTGAATTGTG
 101  CCGTTTTGTC TAAAGTGTTC CCGACCCGTT CGCATACCGT AGCGGCGCAG
 151  GGCGGTATTT CCGCCTCTCT GGGTAATGTG CAGGAAGACC GTTGGGACTG
 201  GCACATGTAC GATACCGTGA AAGGTTCCGA CTGGTTGGGC GACCAAGATG
 251  CGATTGAGTT TATGTGCCGC GCCGCGCCTG AAGCCGTAAT TGAGTTGGAA
 301  CACATGGGTA TGCCTTTTGA CCGTGTGGAA AGCGGTAAAA TTTATCAGCG
 351  TCCTTTCGGC GGCCATACTG CCGAACACGG TAAACGCGCG GTAGAACGCG
 401  CCTGTGCNGT TGCCGACCGT ACAGGTCATG CGATGCTGCA TACTTTGTAC
 451  CAACAAAATG TCCGTGCCAA TACGCAATTC TTTGTGGAAT GGACGGCACA
 501  AGATTTGATT CGTGATGAAA ACGGCGATGT CGTCGGCGTA ACCGCCATGG
 551  AAATGGAAAC CGGCGAAGTT TATATTTTCC ACGCTAAAGC TGTGATGTTT
 601  GCTACCGGCG GCGGCGGCCG TATTTATGCG TCTTCTACCA ATGCCTATAT
 651  GAATACCGGC GATGGTTTGG GTATTTGTGC GCGTGCAGGT ATCCCGTTGG
 701  AAGACATGGA ATTCTGGCAA TTCCACCCGA CCGGCGTGGC AGGTGCGGGC
 751  GTGTTGATTA CCGAAGGCGT ACGCGGCGAG GGCGGTATTC TGTTGAATGC
 801  CGACGGCGAA CGCTTTATGG AACGCTATGC GCCGACCGTA AAAGACTTGG
 851  CTTCTCGCGA CGTTGTTTCC CGCGCGATGG CGATGGAAAT CTACGAAGGT
 901  CGCGGCTGCG GTAAAAACAA AGACCATGTC TTACTGAAAA TCGACCATAT
 951  CGGCGCAGAA AAAATTATGG AAAAACTGCC GGGCATCCGC GAGATTTCCA
1001  TTCAGTTCGC CGGTATCGAT CCGATTAAAG ACCCGATTCC CGTTGTGCCG
1051  ACTACCCACT ATATGATGGG CGGTATTCCG ACCAACTACC ATGGCGAAGT
1101  TGTCGTTCCT CAAGGCGACG AATACGAAGT GCCTGTAAAA GGTCTGTATG
1151  CGGCAGGTGA GTGCGCCTGT GCTTCCGTAC ACGGTGCGAA CCGCTTGGGT
1201  ACGAACTCCC TGCTGGACTT AGTGGTATTC GGTAAAGCTG CCGGCGACAG
1251  CATGATTAAA TTCATCAAAG AGCAAAGCGA CTGGAAACCT TTGCCTGCTA
```

```
1301  ATGCCGGCGA ACTGACCCGC CAACGTATCG AGCGTTTGGA CAATCAAACT
1351  GATGGTGAAA ACGTTGATGC ATTGCGCCGC GAACTGCAAC GCTCCGTACA
1401  ATTGCACGCC GGCGTGTTCC GTACTGATGA GATTCTGAGC AAAGGCGTTC
1451  GAGAAGTCAT GGCGATTGCC GAGCGTGTGA AACGTACCGA AATCAAAGAC
1501  AAGAGCAAAG TGTGGAATAC CGCGCGTATC GAGGCTTTGG AATTGGATAA
1551  CCTAATTGAA GTGGCGAAAG CGACTTTGGT GTCTGCCGAA GCACGTAAAG
1601  AATCACGCGG TGCGCACGCT TCAGACGACC ATCCTGAGCG CGATGATGAA
1651  AACTGGATGA AACATACGCT GTACCATTCA GATGCCAATA CCTTGTCCTA
1701  CAAACCGGTG CACACCAAGC CTTTGAGCGT GGAATACATC AAACCGGCCA
1751  AGCGCGTTTA TTGA
```

This corresponds to the amino acid sequence <SEQ ID 66; ORF 010.a>:

```
a010.pep
  1  MGFPVRKFDA VIVGGGGAGL RAXLQLSKSG LNCAVLSKVF PTRSHTVAAQ
 51  GGISASLGNV QEDRWDWHMY DTVKGSDWLG DQDAIEFMCR AAPEAVIELE
101  HMGMPFDRVE SGKIYQRPFG GHTAEHGKRA VERACAVADR TGHAMLHTLY
151  QQNVRANTQF FVEWTAQDLI RDENGDVVGV TAMEMETGEV YIFHAKAVMF
201  ATGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG
251  VLITEGVRGE GGILLNADGE RFMERYAPTV KDLASRDVVS RAMAMEIYEG
301  RGCGKNKDHV LLKIDHIGAE KIMEKLPGIR EISIQFAGID PIKDPIPVVP
351  TTHYMMGGIP TNYHGEVVVP QGDEYEVPVK GLYAAGECAC ASVHGANRLG
401  TNSLLDLVVF GKAAGDSMIK FIKEQSDWKP LPANAGELTR QRIERLDNQT
451  DGENVDALRR ELQRSVQLHA GVFRTDEILS KGVREVMAIA ERVKRTEIKD
501  KSKVWNTARI EALELDNLIE VAKATLVSAE ARKESRGAHA SDDHPERDDE
551  NWMKHTLYHS DANTLSYKPV HTKPLSVEYI KPAKRVY*
```

m010/a010 98.7% identity over a 231 aa overlap

```
                                        10        20        30
m010.pep                             XQLSKSGLNCAVLSKVFPTRSHTVAAQGGISASXGNV
                                      |||||||||||||||||||||||||||||||| |||
a010            MGFPVRKFDAVIVGGGGAGLRAXLQLSKSGLNCAVLSKVFPTRSHTVAAQGGISASLGNV
                        10        20        30        40        50        60

                 40        50        60        70        80        90
m010.pep        QEDRWDWHMYDTVKGSDWLGDQDAIEFMCRAAPEAVIELEHMGMPFDRVESGKIYQRPFG
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a010            QEDRWDWHMYDTVKGSDWLGDQDAIEFMCRAAPEAVIELEHMGMPFDRVESGKIYQRPFG
                        70        80        90       100       110       120

                100       110       120       130       140       150
m010.pep        GHTAEHGKRAVERXCAVADRTGHAMLHTLYQQNVRANTQFFVEWTAQDLIRDENGDVVGV
                ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
a010            GHTAEHGKRAVERACAVADRTGHAMLHTLYQQNVRANTQFFVEWTAQDLIRDENGDVVGV
                       130       140       150       160       170       180

                160       170       180       190       200       210
m010.pep        TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a010            TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ
                       190       200       210       220       230       240

                220       230
m010.pep        FQPTGVAGAGVLITE
                |:|||||||||||||
a010            FHPTGVAGAGVLITEGVRGEGGILLNADGERFMERYAPTVKDLASRDVVSRAMAMEIYEG
                       250       260       270       280       290       300
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 010 shows 98.7% identity over a 231 aa overlap with a predicted ORF (ORF 010.ng) from *N. gonorrhoeae:*

m010.pep/g010.pep

```
                                                10        20        30
m010.pep                       XQLSKSGLNCAVLSKVFPTRSHTVAAQGGISASXGNV
                                |||||||||||||||||||||||||||||||| |||
g010        MGFPVRKFDAVIVGGGGAGLRAALQLSKSGLNCAVLSKVFPTRSHTVAAQGGISASLGNV
                    10        20        30        40        50        60

             40        50        60        70        80        90
m010.pep    QEDRWDWHMYDTVKGSDWLGDQDAIEFMCRAAPEAVIELEHMGMPFDRVESGKIYQRPFG
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g010        QEDRWDWHMYDTVKGSDWLGDQDAIEFMCRAAPEAVIELEHMGMPFDRVESGKIYQRPFG
                    70        80        90       100       110       120

            100       110       120       130       140       150
m010.pep    GHTAEHGKRAVERXCAVADRTGHAMLHTLYQQNVRANTQFFVEWTAQDLIRDENGDVVGV
            ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
g010        GHTAEHGKRAVERACAVADRTGHAMLHTLYQQNVRANTQFFVEWTAQDLIRDENGDVVGV
                   130       140       150       160       170       180

            160       170       180       190       200       210
m010.pep    TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g010        TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ
                   190       200       210       220       230       240

            220       230
m010.pep    FQPTGVAGAGVLITE
            |:|||||||||||||
g010        FHPTGVAGAGVLITEGVRGEGGILLNADGERFMERYAPTVKDLASRDVVSRAMAMEIYEG
                   250       260       270       280       290       300
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 67>:

```
g010-1.seq..
   1  ATGGGTTTTC CTGTTCGCAA GTTTGATGCC GTGATTGTCG GCGGTGGCGG
  51  TGCAGGTTTA CGTGCAGCCC TCCAATTATC CAAATCCGGT TTGAATTGTG
 101  CCGTTTTGTC TAAAGTGTTC CCGACCCGCT CGCATACCGT AGCGGCGCAG
 151  GGCGGTATTT CCGCCTCTCT GGGTAATGTG CAGGAGGACC GTTGGGACTG
 201  GCACATGTAC GATACCGTGA AAGGTTCCGA CTGGCTGGGC GACCAAGATG
 251  CGATTGAGTT TATGTGTCGC GCTGCGCCTG AAGCGGTGAT TGAGTTGGAA
 301  CACATGGGTA TGCCTTTTGA CCGCGTTGAA AGCGGCAAAA TTTATCAGCG
 351  TCCTTTCGGC GGACATACTG CCGAACATGG TAAACGTGCG GTAGAACGTG
 401  CATGTGCGGT TGCCGACCGT ACCGGTCATG CGATGTTGCA TACTTTGTAC
 451  CAACAAAACG TCCGTGCCAA TACACAATTC TTTGTGGAAT GGACGGCGCA
 501  AGATTTGATT CGTGATGAAA ACGGCGATGT CGTCGGCGTA ACCGCCATGG
 551  AAATGGAAAC GGGCGAAGTT TATATTTTCC ACGCCAAGGC CGTGATGTTT
 601  GCTACCGGTG GCGGCGGTCG TATTTATGCT TCTTCTACCA ATGCTTATAT
 651  GAATACCGGT GACGGTTTGG GCATTTGCGC CCGTGCGGGC ATTCCGTTGG
 701  AAGATATGGA ATTCTGGCAA TTCCACCCGA CCGGCGTGGC GGGTGCGGGC
 751  GTGTTGATTA CCGAAGGCGT ACGCGGCGAG GGCGGTATTC TGTTGAACGC
 801  CGACGGCGAA CGCTTTATGG AACGCTATGC GCCGACCGTA AAAGACTTGG
 851  CTTCTCGCGA CGTGGTTTCA CGCGCGATGG CGATGGAAAT CTATGAAGGT
 901  CGCGGCTGTG GTAAAAACAA AGACCACGTC TTACTGAAAA TCGACCATAT
 951  CGGTGCAGAA AAAATTATGG AAAAACTGCC GGGCATCCGC GAGATTTCCA
1001  TTCAGTTTGC CGGTATCGAT CCGATTAAAG ACCCGATTCC GGTTGTGCCG
1051  ACTACCCACT ATATGATGGG CGGCATTCCG ACCAATTATC ACGGTGAAGT
1101  TGTTGTTCCG CAAGGCGACG AGTACGAAGT ACCTGTAAAA GGCCTGTATG
1151  CCGCAGGTGA GTGCGCCTGT GCTTCCGTAC ACGGTGCGAA CCGTTTGGGT
```

```
1201  ACGAACTCCC TGCTGGACTT GGTGGTGTTC cgcccaaccc cccggtga
```

This corresponds to the amino acid sequence <SEQ ID 68; ORF 010-1.ng>:

```
g010-1.pep
      1  MGFPVRKFDA VIVGGGGAGL RAALQLSKSG LNCAVLSKVF PTRSHTVAAQ
     51  GGISASLGNV QEDRWDWHMY DTVKGSDWLG DQDAIEFMCR AAPEAVIELE
    101  HMGMPFDRVE SGKIYQRPFG GHTAEHGKRA VERACAVADR TGHAMLHTLY
    151  QQNVRANTQF FVEWTAQDLI RDENGDVVGV TAMEMETGEV YIFHAKAVMF
    201  ATGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG
    251  VLITEGVRGE GGILLNADGE RFMERYAPTV KDLASRDVVS RAMAMEIYEG
    301  RGCGKNKDHV LLKIDHIGAE KIMEKLPGIR EISIQFAGID PIKDPIPVVP
    351  TTHYMMGGIP TNYHGEVVVP QGDEYEVPVK GLYAAGECAC ASVHGANRLG
    401  TNSLLDLVVF RPTPR*
```

g010-1 / P10444

```
sp|P10444|DHSA_ECOLI SUCCINATE DEHYDROGENASE FLAVOPROTEIN SUBUNIT
gnl|PID|d1015210 (D90711) Succinate dehydrogenase, flavoprotein  [Escherichia coli] gi|1786942
(AE000175) succinate dehydrogenase flavoprotein subunit [Escherichia coli] Length = 588
 Score = 1073 (495.6 bits), Expect = 6.7e-169, Sum P(2) = 6.7e-169
 Identities = 191/303 (63%), Positives = 238/303 (78%)

Query:     1 MGFPVRKFDAVIVXXXXXXXXXXXXXSKSGLNCAVLSKVFPTRSHTVAAQGGISASLGNV 60
             M  PVR+FDAV++             S+SG  CA+LSKVFPTRS HTV+AQGGI+ +LGN
Sbjct:     1 MKLPVREFDAVVIGAGGAGMRAALQISQSGQTCALLSKVFPTRSHTVSAQGGITVALGNT 60

Query:    61 QEDRWDWHMYDTVKGSDWLGDQDAIEFMCRAAPEAVIELEHMGMPFDRVESGKIYQRPFG 120
              ED W+WHMYDTVKGSD++GDQDAIE+MC+  PEA++ELEHMG+PF R++ G+IYQRPFG
Sbjct:    61 HEDNWEWHMYDTVKGSDYIGDQDAIEYMCKTGPEAILELEHMGLPFSRLDDGRIYQRPFG 120

Query:   121 GHTAEHGKRAVERACAVADRTGHAMLHTLYQQNVRANTQFFVEWTAQDLIRDENGDVVGV 180
             G +   G     R  A ADRTGHA+LHTLYQQN++ +T  F EW A DL+++++G VVG
Sbjct:   121 GQSKNFGGEQAARTAAAADRTGHALLHTLYQQNLKNHTTIFSEWYALDLVKNQDGAVVGC 180

Query:   181 TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ 240
             TA+ +ETGEV  F A+A + ATGG GRIY S+TNA++NTGDG+G+ RAG+P++DME WQ
Sbjct:   181 TALCIETGEVVYFKARATVLATGGAGRIYQSTTNAHINTGDGVGMAIRAGVPVQDMEMWQ 240

Query:   241 FHPTGVAGAGVLITEGVRGEGGILLNADGERFMERYAPTVKDLASRDVVSRAMAMEIYEG 300
             FHPTG+AGAGVL+TEG RGEGG LLN  GERFMERYAP  KDLA RDVV+R++ +EI EG
Sbjct:   241 FHPTGIAGAGVLVTEGCRGEGGYLLNKHGERFMERYAPNAKDLAGRDVVARSIMIEIREG 300

Query:   301 RGC 303
             RGC
Sbjct:   301 RGC 303

 Score = 249 (115.0 bits), Expect = 6.7e-169, Sum P(2) = 6.7e-169
 Identities = 53/102 (51%), Positives = 62/102 (60%)

Query:   309 HVLLKIDHIGAEKIMEKLPGIREISIQFAGXXXXXXXXXXXXXTTHYMMGGIPTNYHGEVV 368
             H  LK+DH+G E +  +LPGI E+S  FA             T HYMMGGIPT   G+ +
Sbjct:   310 HAKLKLDHLGKEVLESRLPGILELSRTFAHVDPVKEPIPVIPTCHYMMGGIPTKVTGQAL 369

Query:   369 VPQGDEYEVPVKGLYAAGECACASVHGANRLGTNSLLDLVVF 410
                   +V V GL+A GE AC SVHGANRLG NSLLDLVVF
Sbjct:   370 TVNEKGEDVVVPGLFAVGEIACVSVHGANRLGGNSLLDLVVF 411
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 69>:

```
m010-1.seq..
      1  ATGGGTTTTC CTGTTCGCAA GTTTGATGCC GTGATTGTCG GCGGTGGTGG
     51  TGCAGGTTTA CGCGCAGCCC TCCAATTATC CAAATCCGGT CTGAATTGTG
    101  CCGTTTTGTC TAAAGTGTTC CCGACCCGTT CGCATACCGT AGCGGCGCAg
    151  GGCGGTATTT CCGCCTCTCT GGGTAATGTG CAGGAAGACC GTTGGGACTG
    201  GCACATGTAC GATACCGTGA AAGGTTCCGA CTGGTTGGGC GACCAAGATG
    251  CGATTGAGTT TATGTGCCGC GCCGCGCCTG AAGCCGTAAT TGAGTTGGAA
    301  CACATGGGTA TGCCTTTTGA CCGTGTGGAA AGCGGTAAAA TTTATCAGCG
    351  TCCTTTCGGC GGCCATACTG CCGAACACGG TAAACGGGCG GTAGAACGCG
    401  CCTGTGCGGT TGCCGACCGT ACAGGTCATG CGATGCTGCA TACTTTGTAC
```

```
 451  CAACAAAACG TCCGTGCCAA TACGCAATTC TTTGTGGAAT GGACGGCACA
 501  AGATTTGATT CGTGATGAAA ACGGCGATGT CGTCGGCGTA ACCGCCATGG
 551  AAATGGAAAC CGGCGAAGTT TATATTTTCC ACGCTAAAGC TGTGATGTTT
 601  GCTACCGGCG GCGGCGGTCG TATTTATGCG TCTTCTACCA ATGCCTATAT
 651  GAATACCGGC GATGGTTTGG GTATTTGTGC GCGTGCAGGT ATCCCGTTGG
 701  AAGACATGGA ATTCTGGCAA TTCCACCCGA CCGGCGTGGC GGGTGCGGGC
 751  GTGTTGATTA CCGAAGGCGT ACGCGGCGAG GGCGGTATTC TGTTGAATGC
 801  CGACGGCGAA CGCTTTATGG AACGCTATGC GCCGACCGTA AAAGACTTGG
 851  CTTCTCGCGA CGTTGTTTCC CGCGCGATGG CGATGGAAAT CTACGAAGGT
 901  CGCGGCTGCG GTAAAAACAA AGACCATGTC TTACTGAAAA TCGACCATAT
 951  CGGCGCAGAA AAAATTATGG AAAAACTGCC GGGCATCCGC GAGATTTCCA
1001  TTCAGTTCGC CGGTATCGAT CCGATTAAAG ACCCGATTCC CGTTGTGCCG
1051  ACTACCCACT ATATGATGGG CGGCATTCCG ACCAATTACC ACGGCGAAGT
1101  TGTCGTTCCG CAAGGTGAAG ATTACGAAGT GCCTGTAAAA GGTCTGTATG
1151  CGGCAGGTGA GTGCGCTTGT GCTTCCGTAC ACGGTGCGAA CCGCTTGGGT
1201  ACCAACTCCC TGTTGGACTT GGTGGTATTC GGTAAAGCTG CCGGCGACAG
1251  CATGATTAAA TTCATCAAAG AGCAAAGCGA CTGGAAACCT TTGCCTGCTA
1301  ATGCAGGTGA GTTGACCCGC CAACGTATCG AGCGTTTGGA CAACCAAACC
1351  GATGGTGAAA ACGTTGATGC ATTGCGTCGC GAACTGCAAC GCTCTGTACA
1401  ACTGCACGCC GGCGTGTTCC GTACTGATGA GATTCTGAGC AAAGGCGTTC
1451  GAGAAGTCAT GGCGATTGCC GAGCGTGTGA AACGTACCGA AATCAAAGAC
1501  AAGAGCAAAG TGTGGAATAC CGCGCGTATC GAGGCTTTGG AATTGGATAA
1551  CCTGATTGAA GTGGCGAAAG CGACTTTGGT GTCTGCCGAA GCACGTAAAG
1601  AATCACGCGG TGCGCACGCT TCAGACGACC ATCCTGAGCG CGATGATGAA
1651  AACTGGATGA AACATACGCT GTACCATTCA GATATCAATA CCTTGTCCTA
1701  CAAACCGGTG CACACCAAGC CTTTGAGCGT GGAATACATC AAACCGGCCA
1751  AGCGCGTTTA TTGATGA
```

This corresponds to the amino acid sequence <SEQ ID 70; ORF 010-1>:

m010-1.pep..

```
  1  MGFPVRKFDA VIVGGGGAGL RAALQLSKSG LNCAVLSKVF PTRSHTVAAQ
 51  GGISASLGNV QEDRWDWHMY DTVKGSDWLG DQDAIEFMCR AAPEAVIELE
101  HMGMPFDRVE SGKIYQRPFG GHTAEHGKRA VERACAVADR TGHAMLHTLY
151  QQNVRANTQF FVEWTAQDLI RDENGDVVGV TAMEMETGEV YIFHAKAVMF
201  ATGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG
251  VLITEGVRGE GGILLNADGE RFMERYAPTV KDLASRDVVS RAMAMEIYEG
301  RGCGKNKDHV LLKIDHIGAE KIMEKLPGIR EISIQFAGID PIKDPIPVVP
351  TTHYMMGGIP TNYHGEVVVP QGEDYEVPVK GLYAAGECAC ASVHGANRLG
401  TNSLLDLVVF GKAAGDSMIK FIKEQSDWKP LPANAGELTR QRIERLDNQT
451  DGENVDALRR ELQRSVQLHA GVFRTDEILS KGVREVMAIA ERVKRTEIKD
501  KSKVWNTARI EALELDNLIE VAKATLVSAE ARKESRGAHA SDDHPERDDE
551  NWMKHTLYHS DINTLSYKPV HTKPLSVEYI KPAKRVY*
```

```
m010-1 / g010-1  99.5% identity in 410 aa overlap

                    10        20        30        40        50        60
m010-1.pep  MGFPVRKFDAVIVGGGGAGLRAALQLSKSGLNCAVLSKVFPTRSHTVAAQGGISASLGNV
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g010-1      MGFPVRKFDAVIVGGGGAGLRAALQLSKSGLNCAVLSKVFPTRSHTVAAQGGISASLGNV
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m010-1.pep  QEDRWDWHMYDTVKGSDWLGDQDAIEFMCRAAPEAVIELEHMGMPFDRVESGKIYQRPFG
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g010-1      QEDRWDWHMYDTVKGSDWLGDQDAIEFMCRAAPEAVIELEHMGMPFDRVESGKIYQRPFG
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m010-1.pep  GHTAEHGKRAVERACAVADRTGHAMLHTLYQQNVRANTQFFVEWTAQDLIRDENGDVVGV
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g010-1      GHTAEHGKRAVERACAVADRTGHAMLHTLYQQNVRANTQFFVEWTAQDLIRDENGDVVGV
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m010-1.pep  TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g010-1      TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ
```

```
                 190        200        210        220        230        240

                 250        260        270        280        290        300
m010-1.pep       FHPTGVAGAGVLITEGVRGEGGILLNADGERFMERYAPTVKDLASRDVVSRAMAMEIYEG
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g010-1           FHPTGVAGAGVLITEGVRGEGGILLNADGERFMERYAPTVKDLASRDVVSRAMAMEIYEG
                 250        260        270        280        290        300

                 310        320        330        340        350        360
m010-1.pep       RGCGKNKDHVLLKIDHIGAEKIMEKLPGIREISIQFAGIDPIKDPIPVVPTTHYMMGGIP
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g010-1           RGCGKNKDHVLLKIDHIGAEKIMEKLPGIREISIQFAGIDPIKDPIPVVPTTHYMMGGIP
                 310        320        330        340        350        360

                 370        380        390        400        410        420
m010-1.pep       TNYHGEVVVPQGEDYEVPVKGLYAAGECACASVHGANRLGTNSLLDLVVFGKAAGDSMIK
                 |||||||||||::|||||||||||||||||||||||||||||||||||||||
g010-1           TNYHGEVVVPQGDEYEVPVKGLYAAGECACASVHGANRLGTNSLLDLVVFRPTPRX
                 370        380        390        400        410

                 430        440        450        460        470        480
m010-1.pep       FIKEQSDWKPLPANAGELTRQRIERLDNQTDGENVDALRRELQRSVQLHAGVFRTDEILS
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 71>:

```
a010-1.seq..
   1  ATGGGCTTTC CTGTTCGCAA GTTTGATGCC GTGATTGTCG GCGGTGGTGG
  51  TGCAGGTTTA CGCGCANCCC TCCAATTATC CAAATCCGGT CTGAATTGTG
 101  CCGTTTTGTC TAAAGTGTTC CCGACCCGTT CGCATACCGT AGCGGCGCAG
 151  GGCGGTATTT CCGCCTCTCT GGGTAATGTG CAGGAAGACC GTTGGGACTG
 201  GCACATGTAC GATACCGTGA AAGGTTCCGA CTGGTTGGGC GACCAAGATG
 251  CGATTGAGTT TATGTGCCGC GCCGCGCCTG AAGCCGTAAT TGAGTTGGAA
 301  CACATGGGTA TGCCTTTTGA CCGTGTGGAA AGCGGTAAAA TTTATCAGCG
 351  TCCTTTCGGC GGCCATACTG CCGAACACGG TAAACGCGCG GTAGAACGCG
 401  CCTGTGCNGT TGCCGACCGT ACAGGTCATG CGATGCTGCA TACTTTGTAC
 451  CAACAAAATG TCCGTGCCAA TACGCAATTC TTTGTGGAAT GGACGGCACA
 501  AGATTTGATT CGTGATGAAA ACGGCGATGT CGTCGGCGTA ACCGCCATGG
 551  AAATGGAAAC CGGCGAAGTT TATATTTTCC ACGCTAAAGC TGTGATGTTT
 601  GCTACCGGCG GCGGCGGCCG TATTTATGCG TCTTCTACCA ATGCCTATAT
 651  GAATACCGGC GATGGTTTGG GTATTTGTGC GCGTGCAGGT ATCCCGTTGG
 701  AAGACATGGA ATTCTGGCAA TTCCACCCGA CCGGCGTGGC AGGTGCGGGC
 751  GTGTTGATTA CCGAAGGCGT ACGCGGCGAG GGCGGTATTC TGTTGAATGC
 801  CGACGGCGAA CGCTTTATGG AACGCTATGC GCCGACCGTA AAAGACTTGG
 851  CTTCTCGCGA CGTTGTTTCC CGCGCGATGG CGATGGAAAT CTACGAAGGT
 901  CGCGGCTGCG GTAAAAACAA AGACCATGTC TTACTGAAAA TCGACCATAT
 951  CGGCGCAGAA AAAATTATGG AAAAACTGCC GGGCATCCGC GAGATTTCCA
1001  TTCAGTTCGC CGGTATCGAT CCGATTAAAG ACCCGATTCC CGTTGTGCCG
1051  ACTACCCACT ATATGATGGG CGGTATTCCG ACCAACTACC ATGGCGAAGT
1101  TGTCGTTCCT CAAGGCGACG AATACGAAGT GCCTGTAAAA GGTCTGTATG
1151  CGGCAGGTGA GTGCGCCTGT GCTTCCGTAC ACGGTGCGAA CCGCTTGGGT
1201  ACGAACTCCC TGCTGGACTT AGTGGTATTC GGTAAAGCTG CCGGCGACAG
1251  CATGATTAAA TTCATCAAAG AGCAAAGCGA CTGGAAACCT TTGCCTGCTA
1301  ATGCCGGCGA ACTGACCCGC CAACGTATCG AGCGTTTGGA CAATCAAACT
1351  GATGGTGAAA ACGTTGATGC ATTGCGCCGC GAACTGCAAC GCTCCGTACA
1401  ATTGCACGCC GGCGTGTTCC GTACTGATGA GATTCTGAGC AAAGGCGTTC
1451  GAGAAGTCAT GGCGATTGCC GAGCGTGTGA AACGTACCGA AATCAAAGAC
1501  AAGAGCAAAG TGTGGAATAC CGCGCGTATC GAGGCTTTGG AATTGGATAA
1551  CCTAATTGAA GTGGCGAAAG CGACTTTGGT GTCTGCCGAA GCACGTAAAG
1601  AATCACGCGG TGCGCACGCT TCAGACGACC ATCCTGAGCG CGATGATGAA
1651  AACTGGATGA AACATACGCT GTACCATTCA GATGCCAATA CCTTGTCCTA
1701  CAAACCGGTG CACACCAAGC CTTTGAGCGT GGAATACATC AAACCGGCCA
1751  AGCGCGTTTA TTGA
```

This corresponds to the amino acid sequence <SEQ ID 72; ORF 010-1.a>:

```
a010-1.pep..
   1  MGFPVRKFDA VIVGGGGAGL RAXLQLSKSG LNCAVLSKVF PTRSHTVAAQ
  51  GGISASLGNV QEDRWDWHMY DTVKGSDWLG DQDAIEFMCR AAPEAVIELE
 101  HMGMPFDRVE SGKIYQRPFG GHTAEHGKRA VERACAVADR TGHAMLHTLY
```

```
151  QQNVRANTQF FVEWTAQDLI RDENGDVVGV TAMEMETGEV YIFHAKAVMF
201  ATGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG
251  VLITEGVRGE GGILLNADGE RFMERYAPTV KDLASRDVVS RAMAMEIYEG
301  RGCGKNKDHV LLKIDHIGAE KIMEKLPGIR EISIQFAGID PIKDPIPVVP
351  TTHYMMGGIP TNYHGEVVVP QGDEYEVPVK GLYAAGECAC ASVHGANRLG
401  TNSLLDLVVF GKAAGDSMIK FIKEQSDWKP LPANAGELTR QRIERLDNQT
451  DGENVDALRR ELQRSVQLHA GVFRTDEILS KGVREVMAIA ERVKRTEIKD
501  KSKVWNTARI EALELDNLIE VAKATLVSAE ARKESRGAHA SDDHPERDDE
551  NWMKHTLYHS DANTLSYKPV HTKPLSVEYI KPAKRVY*
```

m010-1 / a010-1 99.3% identity in 587 aa overlap

```
                    10        20        30        40        50        60
a010-1.pep   MGFPVRKFDAVIVGGGGAGLRAXLQLSKSGLNCAVLSKVFPTRSHTVAAQGGISASLGNV
             ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
a010-1       MGFPVRKFDAVIVGGGGAGLRAALQLSKSGLNCAVLSKVFPTRSHTVAAQGGISASLGNV
                    10        20        30        40        50        60

                    70        80        90       100       110       120
a010-1.pep   QEDRWDWHMYDTVKGSDWLGDQDAIEFMCRAAPEAVIELEHMGMPFDRVESGKIYQRPFG
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m010-1       QEDRWDWHMYDTVKGSDWLGDQDAIEFMCRAAPEAVIELEHMGMPFDRVESGKIYQRPFG
                    70        80        90       100       110       120

                   130       140       150       160       170       180
a010-1.pep   GHTAEHGKRAVERACAVADRTGHAMLHTLYQQNVRANTQFFVEWTAQDLIRDENGDVVGV
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m010-1       GHTAEHGKRAVERACAVADRTGHAMLHTLYQQNVRANTQFFVEWTAQDLIRDENGDVVGV
                   130       140       150       160       170       180

                   190       200       210       220       230       240
a010-1.pep   TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m010-1       TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ
                   190       200       210       220       230       240

                   250       260       270       280       290       300
a010-1.pep   FHPTGVAGAGVLITEGVRGEGGILLNADGERFMERYAPTVKDLASRDVVSRAMAMEIYEG
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m010-1       FHPTGVAGAGVLITEGVRGEGGILLNADGERFMERYAPTVKDLASRDVVSRAMAMEIYEG
                   250       260       270       280       290       300

                   310       320       330       340       350       360
a010-1.pep   RGCGKNKDHVLLKIDHIGAEKIMEKLPGIREISIQFAGIDPIKDPIPVVPTTHYMMGGIP
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m010-1       RGCGKNKDHVLLKIDHIGAEKIMEKLPGIREISIQFAGIDPIKDPIPVVPTTHYMMGGIP
                   310       320       330       340       350       360

                   370       380       390       400       410       420
a010-1.pep   TNYHGEVVVPQGDEYEVPVKGLYAAGECACASVHGANRLGTNSLLDLVVFGKAAGDSMIK
             ||||||||||||::||||||||||||||||||||||||||||||||||||||||||||||
m010-1       TNYHGEVVVPQGEDYEVPVKGLYAAGECACASVHGANRLGTNSLLDLVVFGKAAGDSMIK
                   370       380       390       400       410       420

                   430       440       450       460       470       480
a010-1.pep   FIKEQSDWKPLPANAGELTRQRIERLDNQTDGENVDALRRELQRSVQLHAGVFRTDEILS
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m010-1       FIKEQSDWKPLPANAGELTRQRIERLDNQTDGENVDALRRELQRSVQLHAGVFRTDEILS
                   430       440       450       460       470       480

                   490       500       510       520       530       540
a010-1.pep   KGVREVMAIAERVKRTEIKDKSKVWNTARIEALELDNLIEVAKATLVSAEARKESRGAHA
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m010-1       KGVREVMAIAERVKRTEIKDKSKVWNTARIEALELDNLIEVAKATLVSAEARKESRGAHA
                   490       500       510       520       530       540

                   550       560       570       580
a010-1.pep   SDDHPERDDENWMKHTLYHSDANTLSYKPVHTKPLSVEYIKPAKRVYX
             |||||||||||||||||||| ||||||||||||||||||||||||||
```

```
m010-1          SDDHPERDDENWMKHTLYHSDINTLSYKPVHTKPLSVEYIKPAKRVYX
                           550       560       570       580
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 73>:

g011.seq

```
  1  ATGAAGACAC ACCGCAAGAC CTGCTCTGCG GTGTGTTTTG CTTTTCAGAC
 51  GGCATCGAAA CCCGCCGTTT CCATCCGACA TCCCAGCGAG GACATCATGA
101  GCCTGAAAAC CCGCCTTACC GAAGATATGA AAACCGCGAT GCGCGCCAAA
151  GATCAAGTTT CCCTCGGCAC CATCCGCCTC ATCAATGCCG CCGTCAAACA
201  GTTTGAAGTA GACGAACGCA CCGAAGCCGA CGATGCCAAA ATCACCGCCA
251  TCCTGACCAA AATGGTCAAA CAGCGCAAAG ACGGCGCGAA AATCTACACT
301  GAAGCCGGCC GTCAGGATTT GGCAGACAAA GAAAACGCCG AAATCGACGT
351  GCTGCACCGC TACCTGCCGC AAATGCTCTC CGCCGGCGAA ATCCGCACCG
401  CCGTCGAAGC AGCCGTTGCC GAAACCGGCG CGGCAGGTAT GGCGGATATG
451  GGCAAAGTGA TGGTCGTATT GAAAAcccGC CTCGCCGGCA AAGccgATAT
501  GGGCGAAGTC AACAAAATCT TGAAAAccGt aCTGACCGCC tga
```

This corresponds to the amino acid sequence <SEQ ID 74; ORF 011.ng>:

g011.pep

```
  1  MKTHRKTCSA VCFAFQTASK PAVSIRHPSE DIMSLKTRLT EDMKTAMRAK
 51  DQVSLGTIRL INAAVKQFEV DERTEADDAK ITAILTKMVK QRKDGAKIYT
101  EAGRQDLADK ENAEIDVLHR YLPQMLSAGE IRTAVEAAVA ETGAAGMADM
151  GKVMVVLKTR LAGKADMGEV NKILKTVLTA *
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 75>:

m011.seq (partial)

```
  1  ATGAGGACAC ACCGCAAGAC CTGCTCTGCG GTGTGTTTTG CTTTTCAGAC
 51  GGCATCGAAA CCCGCCGTTT CCATCCGACA TCCCAGCGAG GACATCATGA
101  GCCTGAAAAT CCGCCTTACC GAAGACATGA AAACCGCGAT GCGCGCCAAA
151  GACCAAGTTT CCCTCGGCAC CATCCGCCTC ATCAACGCCG CCGTCAAACA
201  GTTTGAAGTG GACGAACGCA CCGAAGCCGA CGATGCCAAA ATCACCGCCA
251  TCCTGACCAA AATGGTCAAA CAGCGAAAAG ACAGCGCGAA AATCTACACT
301  GAAGCCGGCC GTCAGGATTT GGCAGACAAA GAAAACGCCG AAATCGAGGT
351  ACTGCACCGC TACCTTCCCC AAATGCTTTC CGCCGGCGAA ATCCGTACCG
401  AGGTCGAAGC TGCCGTTGCC GAAACCGGCG CGGCAGGTAT GGCGGATATG
451  GGTAAAGTCA TGGGGCTGCT GAAAACCCGC CTCGCAGGTA AAGCCGA...
```

This corresponds to the amino acid sequence <SEQ ID 76; ORF 011>:

m011.pep (partial)

```
  1  MRTHRKTCSA VCFAFQTASK PAVSIRHPSE DIMSLKIRLT EDMKTAMRAK
 51  DQVSLGTIRL INAAVKQFEV DERTEADDAK ITAILTKMVK QRKDSAKIYT
101  EAGRQDLADK ENAEIEVLHR YLPQMLSAGE IRTEVEAAVA ETGAAGMADM
151  GKVMGLLKTR LAGKA.....
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 011 shows 95.8% identity over a 165 aa overlap with a predicted ORF (ORF 011.ng) from *N. gonorrhoeae:*

m011/g011

```
                      10        20        30        40        50        60
m011.pep     MRTHRKTCSAVCFAFQTASKPAVSIRHPSEDIMSLKIRLTEDMKTAMRAKDQVSLGTIRL
             |:|||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
g011         MKTHRKTCSAVCFAFQTASKPAVSIRHPSEDIMSLKTRLTEDMKTAMRAKDQVSLGTIRL
                      10        20        30        40        50        60

                      70        80        90       100       110       120
m011.pep     INAAVKQFEVDERTEADDAKITAILTKMVKQRKDSAKIYTEAGRQDLADKENAEIEVLHR
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||:||||
g011         INAAVKQFEVDERTEADDAKITAILTKMVKQRKDGAKIYTEAGRQDLADKENAEIDVLHR
```

```
                    70        80        90       100       110       120

                   130       140       150       160
m011.pep    YLPQMLSAGEIRTEVEAAVAETGAAGMADMGKVMGLLKTRLAGKA
            ||||||||||||| ||||||||||||||||||||:| :|||||||||
g011        YLPQMLSAGEIRTAVEAAVAETGAAGMADMGKVMVVLKTRLAGKADMGEVNKILKTVLTA
                   130       140       150       160       170       180

g011        X
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 77>:

```
g012.seq
      1  ATGCTCGCCC GTCGCTATTT TTTCAATATC CAACCCGGGG CGGTTTTCAC
     51  TGACAAACTG CTTGAACAAC TGATGCGTTT CCTCCAGTTC CTGCCGGAAT
    101  TTCTGTTTGC CCTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCGCTT
    151  AAATTCGCCC GCCGTCATCA CATCCACATC AATATCATGT TTTTTCAACa
    201  gGcggTGGAT ATTCGgcact tccgCcacca cacccaccga accgatgacc
    251  gcaaacggaG CGGAAACAAT TTTATCCGCc acacacgcca tcatatagcc
    301  gcCGCTTGCC GCGACCTTAT CGAcggcgac ggTCAGCGGA ATATTGCGTT
    351  CGCGCAAACG CCTAAGCTGC GAAGCCGCCA AACCGTAACC GTGAACCACG
    401  CCGCCCGGAC TTTCCAATCT GAGCAGAACC TCATCTTCAG GCTTGGCAAT
    451  CAAAAGCACC GCCGTAATCT CATGACGCAA GGATTCTACG GCGTGTGCAT
    501  ACAAATCGCC GTCAAAATCC AACACAAAAA GGCGGGATTT TTGCGTTTCG
    551  GCAGATTTCT CCCCGCCCTC CTTCAAACGC TTTTTCTCTG CTTTGGCTTC
    601  CGCCTTTTCC TTTTTCTTTT CTTTTTTTTC CTGATGTTTT GTCTCTTCCT
    651  CGCTTAA
```

This corresponds to the amino acid sequence <SEQ ID 78; ORF 012.ng>:

```
g012.pep
      1  MLARRYFFNI QPGAVFTDKL LEQLMRFLQF LPEFLFALFR IFTHKSNRAL
     51  KFARRHHIHI NIMFFQQAVD IRHFRHHTHR TDDRKRSGNN FIRHTRHHIA
    101  AACRDLIDGD GQRNIAFAQT PKLRSRQTVT VNHAARTFQS EQNLIFRLGN
    151  QKHRRNLMTQ GFYGVCIQIA VKIQHKKAGF LRFGRFLPAL LQTLFLCFGF
    201  RLFLFLFFFF LMFCLFLA*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 79>:

```
m012.seq
      1  ATGCTCGCCC GTTGCCACTT CCTCAATATC CAATTGAGGG CGGTTCTCGC
     51  TGACAAACTG CTTGAACAAC TGATGCGTTT CCTCCAGTTC CTGTCGGAAT
    101  TTCTGTTTGC CCTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCGCTT
    151  AAATTCGCCC GCCGTCATCA CATCCACATC AATATCATGT TTTTTCAACA
    201  GGCGGTGGAT ATTCGGTACT TCCGCCACCA CACCCACCGA ACCGACAATC
    251  GCAAACGGAG CGGAAGCAAT TTTATCCGCC ACACACGCCA TCATATAACC
    301  GCCGCTCGCn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
    351  nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
    401  nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
    451  nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
    501  nnnnnnnnnn nnnnnnnnnC AACACAAAAA GGCGTGATTT nTGCGTTTCG
    551  GCAGATTTCT CCCCACCCTC CTTCAAACGT TTTTCcTCTG CTTTGGCTTC
    601  CGCCTTTTCC TTTTTCTTTT CCTCTTTTTC CTGATGTTGT GCCTCTTCCC
    651  CGCTTAA
```

This corresponds to the amino acid sequence <SEQ ID 80; ORF 012>:

```
m012.pep
      1  MLARCHFLNI QLRAVLADKL LEQLMRFLQF LSEFLFALFR IFTHKSNRAL
     51  KFARRHHIHI NIMFFQQAVD IRYFRHHTHR TDNRKRSGSN FIRHTRHHIT
    101  AARXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX
    151  XXXXXXXXXX XXXXXXXXXX XXXQHKKA*F XRFGRFLPTL LQTFFLCFGF
    201  RLFLFLFLFF LMLCLFPA*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 81>:

```
a012.seq
      1  ATGCTCGCCC GTTGCCACTT CCTCAATATC CAATTGAGGG CGGTTCTCGC
     51  TGACAAACTG CTTGAACAAC TGATGCGTTT CCTCCAGTTC CTGTCGGAAT
    101  TTCTGTTTGC CCTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCGCTT
    151  AAATTCGCCC GCCGTCATCA CATCCACATC AATATCATGT TTTTTCAACA
    201  GGCGGTGGAT ATTCGGTACT TCCGCTACAA CACCCACCGA ACCGACAATC
    251  GCAAACGGAG CGGAAACAAT TTTATCCGCC ACACACGCCA TCATATAACC
    301  ACCGCTCGCC GCCACCTTAT CGACGGCGAC GGTCAGCGGA ATATTGCGTT
    351  CGCGCAAACG CCTAAGCTGC GAAGCCGCCA AACCGTAACC GTGAACCACG
    401  CCGCCCGGAC TTTCCAATCT AAGCAGAACC TCATCTTCAG GCTTGGCAAT
    451  CAAAAGCACC GCCGTAATCT CATGACGCAA GGATTCTACG GCGTGTGCAT
    501  ACAAATCGCC GTCAAAATCC AACACAAAAA GGCGGGATTT TTGCGTTTCG
    551  GAAGATTTCT CCCCACCCTC CTTCAAACGC TTTTTCTCTG CTTTGGCTTC
    601  CGCCTTTTCC TTTTTCTTTT CCTCTTTTTC CTGATGTTTT GCCTCTTCCC
    651  CGCTTAA
```

This corresponds to the amino acid sequence <SEQ ID 82; ORF 012.a>:

```
a012.pep
      1  MLARCHFLNI QLRAVLADKL LEQLMRFLQF LSEFLFALFR IFTHKSNRAL
     51  KFARRHHIHI NIMFFQQAVD IRYFRYNTHR TDNRKRSGNN FIRHTRHHIT
    101  TARRHLIDGD GQRNIAFAQT PKLRSRQTVT VNHAARTFQS KQNLIFRLGN
    151  QKHRRNLMTQ GFYGVCIQIA VKIQHKKAGF LRFGRFLPTL LQTLFLCFGF
    201  RLFLFLFLFF LMFCLFPA*
```

m012/a012 64.2% identity over a 218 aa overlap

```
                    10        20        30        40        50        60
m012.pep    MLARCHFLNIQLRAVLADKLLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a012        MLARCHFLNIQLRAVLADKLLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m012.pep    NIMFFQQAVDIRYFRHHTHRTDNRKRSGSNFIRHTRHHITAARXXXXXXXXXXXXXXXXX
            |||||||||||||||::|||||||||||:|||||||||||:||                :
a012        NIMFFQQAVDIRYFRYNTHRTDNRKRSGNNFIRHTRHHITTARRHLIDGDGQRNIAFAQT
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m012.pep    XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXQHKKAXF
             : :              :                       :          ||||| |
a012        PKLRSRQTVTVNHAARTFQSKQNLIFRLGNQKHRRNLMTQGFYGVCIQIAVKIQHKKAGF
                   130       140       150       160       170       180

                   190       200       210      219
m012.pep    XRFGRFLPTLLQTFFLCFGFRLFLFLFLFFLMLCLFPAX
             |||||||||||||:||||||||||||||||||:|||||
a012        LRFGRFLPTLLQTLFLCFGFRLFLFLFLFFLMFCLFPAX
                   190       200       210
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 012 shows 58.7% identity over a 218 aa overlap with a predicted ORF (ORF 012.ng) from *N. gonorrhoeae*:

m012/g012

```
                    10        20        30        40        50        60
m012.pep    MLARCHFLNIQLRAVLADKLLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI
            ||||  :|:|||  ||::|||||||||||||| ||||||||||||||||||||||||||||
g012        MLARRYFFNIQPGAVFTDKLLEQLMRFLQFLPEFLFALFRIFTHKSNRALKFARRHHIHI
```

```
                    10        20        30        40        50        60
                    70        80        90       100       110       120
m012.pep    NIMFFQQAVDIRYFRHHTHRTDNRKRSGSNFIRHTRHHITAARXXXXXXXXXXXXXXXXX
            ||||||||||||:|||||||||:|||||:||||||||||:||                 :
g012        NIMFFQQAVDIRHFRHHTHRTDDRKRSGNNFIRHTRHHIAAACRDLIDGDGQRNIAFAQT
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m012.pep    XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXQHKKAXF
                   : :      :                     :                 ||||| |
g012        PKLRSRQTVTVNHAARTFQSEQNLIFRLGNQKHPRNLMTQGFYGVCIQIAVKIQHKKAGF
                   130       140       150       160       170       180

                   190       200       210      219
m012.pep    XRFGRFLPTLLQTFFLCFGFRLFLFLFLFFLMLCLFPAX
             |||||||:||||:|||||||||||||:||||:||| ||
g012        LRFGRFLPALLQTLFLCFGFRLFLFLFFFFLMFCLFLAX
                   190       200       210
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 83>:

```
m012-1.seq
      1  ATGCTCGCCC GTTGCCACTT CCTCAATATC CAATTGAGGG CGGTTCTCGC
     51  TGACAAACTG CTTGAACAAC TGATGCGTTT CCTCCAGTTC CTGTCGGAAT
    101  TTCTGTTTGC CCTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCGCTT
    151  AAATTCGCCC GCCGTCATCA CATCCACATC AATATCATGT TTTTTCAACA
    201  GGCGGTGGAT ATTCGGTACT TCCGCCACCA CACCCACCGA ACCGACAATC
    251  GCAAACGGAG CGGAAGCAAT TTTATCCGCC ACACACGCCA TCATATAACC
    301  GCCGCTCGCC GCCACCTTAT CGACGGCGAC GGTCAGCGGA ATATTGCGTT
    351  CGCGCAAACG CyTAAGCTGC GAAGCCGCCA AACCGTAACC GTGAACCACG
    401  CCGCCCGGAC TTTCCAATCT GAGCAGAACC TCATCTTCAG GCTTGGCAAT
    451  CAAAAGCACC GCCGTAATCT CATGACGCAA GGATTCTACG GCGTGTGCAT
    501  ACAAATCGCC GTCAAAATCC AACACAAAAA GGCGGGATTT TTGCGTTTCG
    551  GCAGATTTCT CCCCACCCTC CTTCAAACGC TTTTTCTCTG CTTTGGCTTC
    601  CGCCTTTTCC TTTTTCTTTT CCTCTTTTTC CTGATGTTTT GCCTCTTCCC
    651  CGCTTAA
```

This corresponds to the amino acid sequence <SEQ ID 84; ORF 012-1>:

```
m012-1.pep
      1  MLARCHFLNI QLRAVLADKL LEQLMRFLQF LSEFLFALFR IFTHKSNRAL
     51  KFARRHHIHI NIMFFQQAVD IRYFRHHTHR TDNRKRSGSN FIRHTRHHIT
    101  AARRHLIDGD GQRNIAFAQT XKLRSRQTVT VNHAARTFQS EQNLIFRLGN
    151  QKHRRNLMTQ GFYGVCIQIA VKIQHKKAGF LRFGRFLPTL LQTLFLCFGF
    201  RLFLFLFLFF LMFCLFPA*
```

m012-1/g012 91.7% identity in 218 aa overlap

```
                    10        20        30        40        50        60
m012-1.pep  MLARCHFLNIQLRAVLADKLLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI
            |||| :|:|||  ||::|||||||||||||| ||||||||||||||||||||||||||||
g012        MLARRYFFNIQPGAVFTDKLLEQLMRFLQFLPEFLFALFRIFTHKSNRALKFARRHHIHI
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m012-1.pep  NIMFFQQAVDIRYFRHHTHRTDNRKRSGSNFIRHTRHHITAARRHLIDGDGQRNIAFAQT
            ||||||||||||:|||||||||:|||||:||||||||||:|| | |||||||||||||||
g012        NIMFFQQAVDIRHFRHHTHRTDDRKRSGNNFIRHTRHHIAAACRDLIDGDGQRNIAFAQT
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m012-1.pep  XKLRSRQTVTVNHAARTFQSEQNLIFRLGNQKHRRNLMTQGFYGVCIQIAVKIQHKKAGF
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g012        PKLRSRQTVTVNHAARTFQSEQNLIFRLGNQKHRRNLMTQGFYGVCIQIAVKIQHKKAGF
```

```
                   130       140       150       160       170       180

                   190       200       210      219
m012-1.pep    LRFGRFLPTLLQTLFLCFGFRLFLFLFLFFLMFCLFPAX
              |||||||:|||||||||||||||||||:|||||||| ||
g012          LRFGRFLPALLQTLFLCFGFRLFLFLFFFFLMFCLFLAX
                   190       200       210
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 85>:

a012-1.seq

```
  1  ATGCTCGCCC GTTGCCACTT CCTCAATATC CAATTGAGGG CGGTTCTCGC
 51  TGACAAACTG CTTGAACAAC TGATGCGTTT CCTCCAGTTC CTGTCGGAAT
101  TTCTGTTTGC CCTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCGCTT
151  AAATTCGCCC GCCGTCATCA CATCCACATC AATATCATGT TTTTTCAACA
201  GGCGGTGGAT ATTCGGTACT TCCGCTACAA CACCCACCGA ACCGACAATC
251  GCAAACGGAG CGGAAACAAT TTTATCCGCC ACACACGCCA TCATATAACC
301  ACCGCTCGCC GCCACCTTAT CGACGGCGAC GGTCAGCGGA ATATTGCGTT
351  CGCGCAAACG CCTAAGCTGC GAAGCCGCCA AACCGTAACC GTGAACCACG
401  CCGCCCGGAC TTTCCAATCT AAGCAGAACC TCATCTTCAG GCTTGGCAAT
451  CAAAAGCACC GCCGTAATCT CATGACGCAA GGATTCTACG GCGTGTGCAT
501  ACAAATCGCC GTCAAAATCC AACACAAAAA GGCGGGATTT TTGCGTTTCG
551  GAAGATTTCT CCCCACCCTC CTTCAAACGC [illegible]TTTCTCTG CTTTGGCTTC
601  CGCCTTTTCC TTTTTCTTTT CCTCTTTTTC CTGATGTTTT GCCTCTTCCC
651  CGCTTAA
```

This corresponds to the amino acid sequence <SEQ ID 86; ORF 012-1.a>:

```
a012-1.pep
  1  MLARCHFLNI QLRAVLADKL LEQLMRFLQF LSEFLFALFR IFTHKSNRAL
 51  KFARRHHIHI NIMFFQQAVD IRYFRYNTHR TDNRKRSGNN FIRHTRHHIT
101  TARRHLIDGD GQRNIAFAQT PKLRSRQTVT VNHAARTFQS KQNLIFRLGN
151  QKHRRNLMTQ GFYGVCIQIA VKIQHKKAGF LRFGRFLPTL LQTLFLCFGF
201  RLFLFLFLFF LMFCLFPA*
```

a012-1/m012-1   97.2% identity in 218 aa overlap

```
                   10        20        30        40        50        60
a012-1.pep    MLARCHFLNIQLRAVLADKLLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m012-1        MLARCHFLNIQLRAVLADKLLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI
                   10        20        30        40        50        60

                   70        80        90       100       110       120
a012-1.pep    NIMFFQQAVDIRYFRYNTHRTDNRKRSGNNFIRHTRHHITTARRHLIDGDGQRNIAFAQT
              |||||||||||||||::||||||||||||:||||||||||:|||||||||||||||||||
m012-1        NIMFFQQAVDIRYFRHHTHRTDNRKRSGSNFIRHTRHHITAARRHLIDGDGQRNIAFAQT
                   70        80        90       100       110       120

                  130       140       150       160       170       180
a012-1.pep    PKLRSRQTVTVNHAARTFQSKQNLIFRLGNQKHRRNLMTQGFYGVCIQIAVKIQHKKAGF
               |||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
m012-1        XKLRSRQTVTVNHAARTFQSEQNLIFRLGNQKHRRNLMTQGFYGVCIQIAVKIQHKKAGF
                  130       140       150       160       170       180

                  190       200       210      219
a012-1.pep    LRFGRFLPTLLQTLFLCFGFRLFLFLFLFFLMFCLFPAX
              |||||||||||||||||||||||||||||||||||||||
m012-1        LRFGRFLPTLLQTLFLCFGFRLFLFLFLFFLMFCLFPAX
                  190       200       210
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 87>:

g013.seq

```
  1  aTgcctttga ccatgctgtg cagcaGGAcg tGCGGTTtgt tcataataca
 51  gtCcgaccGG AAAagcggAG GAAaCGCAGT GCCGC[illegible]CCT TCCCCTTTCT
101  TGCCGTGGCA GGCGATGCag tTgGATTCGT ACACTTTTTG CCCTTTtGtc
```

```
151  atgatGCTgt tgtcggCGGC AGAAGCgGCG GcgCAGAGGC AGCACAAGAT
201  GAAGGCGGTC GGCAGTCGGG TTGTGTtcat tGgcgTTTCC cctaatgttt
251  tgaaaccttg ttttttgatt Ttgcctttac ggggtgaaaa gtttttTtgg
301  cccaaatccg gaatttag
```

This corresponds to the amino acid sequence <SEQ ID 88; ORF 013.ng:

g013.pep

```
  1  MPLTMLCSRT CGLFIIQSDR KSGGNAVPRP SPFLPWQAMQ LDSYTFCPFV
 51  MMLLSAAEAA AQRQHKMKAV GSRVVFIGVS PNVLKPCFLI LPLRGEKFFW
101  PKSGI*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 89>:

m013.seq

```
  1  ATGCCTTTGA CCATGCTGTG CAGCAGCACC TGCGGTTTTT TCATGATGAA
 51  GTCGGAGCGG TAGAGCGGCG GAAACATGGT TCCGCGGCCT TCGCCCTTTT
101  TGCCGTGGCA GGCGACGCAG TTGGATTCGT ACACTTTTTG CCCTTTTGTC
151  ATGATGCTGT TGTCGGCGGC AGAAGCGGCG GCGCAGAAGC AGCCCAAGAC
201  GAGGGCGGTC GGCAGTCGGG TTGTGTTCAT TGGTGTTTCC TTCATGTTTG
251  AAACCTTGTT GTTGATTTTG CGTAGCGGGT GAAAGATTTT TTTGCCGAAT
301  CAGTAG
```

This corresponds to the amino acid sequence <SEQ ID 90; ORF 013>:

m013.pep

```
  1  MPLTMLCSST CGFFMMKSER XSGGNMVPRP SPFLPWQATQ LDSYTFCPFV
 51  MMLLSAAEAA AQKQPKTRAV GSRVVFIGVS FMFETLLLIL RSGXKIFLPN
101  Q*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 91>:

a013.seq

```
  1  ATGCCTTTGA CCATGCTGTG CAGCAGCACC TGCGGTTTTT TCATGATGAA
 51  GTCGGAGCGG TAGAGCGGCG GAAACATGGT TCCGCGGCCT TCGCCCTTTT
101  TGCCGTGGCA GGCGACGCAG TTGGATTCGT ACACTTTTTG CCCTTTTGTC
151  ATGATGCTGT TGTCGGCGGC AGAAGCGGCG GCGCAGAGGC AGCCCAAGAC
201  GAGGGCGGTC GGCAGTCGGG TTGTGTTCAT TGGTGTTTCC TTAATGTTTG
251  AAACCTTGTT GTTGATTTTG CGTAGCGGGT GAAAGATTTT CTTGCCGAAT
301  CGGTAG
```

This corresponds to the amino acid sequence <SEQ ID 92; ORF 013.a>:

a013.pep

```
  1  MPLTMLCSST CGFFMMKSER *SGGNMVPRP SPFLPWQATQ LDSYTFCPFV
 51  MMLLSAAEAA AQRQPKTRAV GSRVVFIGVS LMFETLLLIL RSG*KIFLPN
101  R*
```

m013/a013 97.0% identity over a 101 aa overlap

```
                    10        20        30        40        50        60
m013.pep    MPLTMLCSSTCGFFMMKSERXSGGNMVPRPSPFLPWQATQLDSYTFCPFVMMLLSAAEAA
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a013        MPLTMLCSSTCGFFMMKSERXSGGNMVPRPSPFLPWQATQLDSYTFCPFVMMLLSAAEAA
                    10        20        30        40        50        60

                    70        80        90       100
m013.pep    AQKQPKTRAVGSRVVFIGVSFMFETLLLILRSGXKIFLPNQX
            ||:|||||||||||||||||:|||||||||||||||||||:|
a013        AQRQPKTRAVGSRVVFIGVSLMFETLLLILRSGXKIFLPNRX
                    70        80        90       100
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 013 shows 73.3% identity over a 101 aa overlap with a predicted ORF (ORF 013.ng) from *N. gonorrhoeae:*

m013/g013

```
                   10        20        30        40        50        60
m013.pep   MPLTMLCSSTCGFFMMKSERXSGGNMVPRPSPFLPWQATQLDSYTFCPFVMMLLSAAEAA
           |||||||| |||:|:::|:| |||| |||||||||||| |||||||||||||||||||||
g013       MPLTMLCSRTCGLFIIQSDRKSGGNAVPRPSPFLPWQAMQLDSYTFCPFVMMLLSAAEAA
                   10        20        30        40        50        60

                   70        80         90        100
m013.pep   AQKQPKTRAVGSRVVFIGVSF-MFETLLLILR-SGXKIFLPNQX
           ||:| | :||||||||||||  :::  :|||   | |:| |:
g013       AQRQHKMKAVGSRVVFIGVSPNVLKPCFLILPLRGEKFFWPKSGIX
                   70        80        90       100
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 93>:

```
g015.seq
      1  ATGCAGTATC TGATTGTCAA ATACAGCCAT CAAATCTTCG TTACCATCAC
     51  CATTTTGGTA TTCAACATCC GTTTTTTCCT ACTTTGGAAA AATCCAGAAA
    101  AGCCCTTGGT CGGCTTTTGG AAAGCACTGC CCCACCTCAA CGACACGATG
    151  CTGCTGTTTA CGGGATTGTG GCTGATGAAG ATTACCCATT TCTCCCCGTT
    201  CAACGCGCCT TGGCTCGGCA CAAAAATCCT GCTCCTGTTC GCCTACATCG
    251  CACTGGGCAT GGTAATGATG CGCGCCCGTC CGCGTTCGAC CAAGTTCTAC
    301  ACCGTTTACC TGCTCGCTAT GTGTTGCATC GCCTGCATCG TTTACCTTGC
    351  CAAAACCAAA GTCCTGCCAT TCTGA
```

This corresponds to the amino acid sequence <SEQ ID 94; ORF 015.ng>:

```
g015.pep
      1  MQYLIVKYSH QIFVTITILV FNIRFFLLWK NPEKPLVGFW KALPHLNDTM
     51  LLFTGLWLMK ITHFSPFNAP WLGTKILLLF AYIALGMVMM RARPRSTKFY
    101  TVYLLAMCCI ACIVYLAKTK VLPF*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 95>:

```
m015.seq   (partial)
      1  ..AAAATCAGAA AAGCCTTGGC GGGCTTTTGG AAGGCACTGC CCCACCTTAA
     51    CGACACCATG CTGCTGTTTA CGGGATTGTG GCTGATGAAA ATTACCCATT
    101    TCTCCCCGTT CAACGCGCCT TGGCTCGGTA CAAAAATCCT GCTTCTGCTC
    151    GCCTATATCG CATTGGGTAT GATGATGATG CGCGCCCGTC CGCGTTCGAC
    201    CAAGTTCTAC ACCGTTTACC TGCTCGCCAT GTGTTGCGTC GCCTGCATCG
    251    TTTACCTTGC CAAAACCAAA GTCCTGCCTT TCTGA
```

This corresponds to the amino acid sequence <SEQ ID 96; ORF 015:

```
m015.pep  (partial)
      1  ..KIRKALAGFW KALPHLNDTM LLFTGLWLMK ITHFSPFNAP WLGTKILLLL
     51    AYIALGMMMM RARPRSTKFY TVYLLAMCCV ACIVYLAKTK VLPF*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 97>:

```
a015.seq
      1  ATGCAGTATC TGATTGTCAA ATACAGCCAT CAAATCTTCG TTACCATCAC
     51  CATTTTGGTA TTCAACATCC GTGTTTTCNT ACTTTGGAAA AATCCAGAAA
    101  AGCCCTTGGC GGGCTTTTGG AAGGCACTGC CCCACCTTAA CGACACCATG
    151  CTGCTGTTTA CGGGATTGTG GCTGATGAAA ATTACCCATT TCTCCCCGTT
    201  CAACGCGCCT TGGCTCGGTA CAAAAATCCT GCTTCTGCTC GCCTATATCG
    251  CATTGGGTAT GATGATGATG CGCGCCCGTC CGCGTTCGAC CAAGTTCTAC
    301  ACCGTTTACC TGCTCGCCAT GTGTTGCCTC ACCTGCATCG TTTACCTTGC
    351  CAAAACCAAA GTCCTGCCTT TCTGA
```

This corresponds to the amino acid sequence <SEQ ID 98; ORF 015.a>:

```
a015.pep
```

```
  1  MQYLIVKYSH QIFVTITILV FNIRVFXLWK NPEKPLAGFW KALPHLNDTM
 51  LLFTGLWLMK ITHFSPFNAP WLGTKILLLL AYIALGMMMM RARPRSTKFY
101  TVYLLAMCCL TCIVYLAKTK VLPF*
```

m015/a015 96.7% identity over a 91 aa overlap

```
                                                 10        20        30
m015.pep                                 KIRKALAGFWKALPHLNDTMLLFTGLWLMKITH
                                           | |||||||||||||||||||||||||||||||
a015         LIVKYSHQIFVTITILVFNIRVFXLWKNPEKPLAGFWKALPHLNDTMLLFTGLWLMKITH
                     10        20        30        40        50        60

                     40        50        60        70        80        90
m015.pep     FSPFNAPWLGTKILLLLAYIALGMMMMRARPRSTKFYTVYLLAMCCVACIVYLAKTKVLP
             ||||||||||||||||||||||||||||||||||||||||||||||||::||||||||||
a015         FSPFNAPWLGTKILLLLAYIALGMMMMRARPRSTKFYTVYLLAMCCLTCIVYLAKTKVLP
                     70        80        90       100       110       120

m015.pep     FX
             ||
a015         FX
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 015 shows 94.5% identity over a 91 aa overlap with a predicted ORF (ORF 015.ng) from *N. gonorrhoeae:*

```
m015/g015

                                                 10        20        30
m015.pep                                 KIRKALAGFWKALPHLNDTMLLFTGLWLMKITH
                                           | |:|||||||||||||||||||||||||||||
g015         LIVKYSHQIFVTITILVFNIRFFLLWKNPEKPLVGFWKALPHLNDTMLLFTGLWLMKITH
                     10        20        30        40        50        60

                     40        50        60        70        80        90
m015.pep     FSPFNAPWLGTKILLLLAYIALGMMMMRARPRSTKFYTVYLLAMCCVACIVYLAKTKVLP
             ||||||||||||||||:|||||||:|||||||||||||||||||||:|||||||||||||
g015         FSPFNAPWLGTKILLLFAYIALGMVMMRARPRSTKFYTVYLLAMCCIACIVYLAKTKVLP
                     70        80        90       100       110       120

m015.pep     FX
             ||
g015         FX
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 99>:

```
g018.seq
  1  atGCAGCAGG GGCagttggt tggacgcgtc gcccgcaata AAGATATGCG
 51  GAATgctggt CTGCATggtC AGCGGATCGG CAACGGGtac gccgcgcgcg
101  tctttgTCGA TATTGATGTT TTCCAAACCG ATATtgTCAA CGTTCGGACG
151  GCgACCTACG GCTGCCAACA TATATTCGGC AACAAATACG CCTTTTTCGC
201  CATCCTGCTC CCAATGGACT tctACATTGC CGTCTGCGTC GAGTTTGACC
251  TCGGTTTTAG CATCCAGATG CAGTTTCAAT tctTCTCCGA ACACGGCTTT
301  CGCCTCGTCT GA
```

This corresponds to the amino acid sequence <SEQ ID 100; ORF 018.ng>:

```
g018.pep
  1  MQQGQLVGRV ARNKDMRNAG LHGQRIGNGY AARVFVDIDV FQTDIVNVRT
 51  ATYGCQHIFG NKYAFFAILL PMDFYIAVCV EFDLGFSIQM QFQFFSEHGF
```

```
101  RLV*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 101>:

m018.seq

```
  1  ATGCAGCAGA GGCAGTTGGT TGGACGCATC GCCTGCGATG AAGATATGCG
 51  GAATACTGGT CTGCATGGTC AGCGGGTCGG CAACAGGTAC GCCGCGCGCA
101  TCTTTTTCGA TATTGATATT TTCCAAACCG ATATTGTCAA CGTTCGGACG
151  GCGGCCCACG GCTGCCAGCA TATATTCGGC AACAAATACG CCTTTTTCGC
201  CATCCTGCTC CCAATGGACT TCTACATTGC CGTCTGCATC GAGTTTGACC
251  TCGGTTTTAG CATCCAGATG CAGTTTCAAT TCTTCGCCGA ACACGGCGTT
301  CGCCTCGTCT GA
```

This corresponds to the amino acid sequence <SEQ ID 102; ORF 018>:

m018.pep

```
  1  MQQRQLVGRI ACDEDMRNTG LHGQRVGNRY AARIFFDIDI FQTDIVNVRT
 51  AAHGCQHIFG NKYAFFAILL PMDFYIAVCI EFDLGFSIQM QFQFFAEHGV
101  RLV*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 103>:

a018.seq

```
  1  ATGCAGCAGG GGCAGTTGGT TGGACGCGTC GCCCGCAATA AAGATATGCG
 51  GAATACTGGT CTGCATAGTC AGCGGATCGG CAACGGGTAC GCCGCGCGCA
101  TCTTTTTCGA TATTGATGTT TTCCAAACCG ATATTGTCAA CGTTCGGACG
151  GCGGCCTACG GCTGCCAGCA TATATTCGGC AACAAATACG CCTTTTTCGC
201  CATCCTGCTC CCAATGGACT TCTACATTGC CGTCTGCGTC GAGTTTGGCC
251  TCGGTTTTAG CATCCAAATG CAGTTTCAAT TCTTCACCGA ACACGGCTTT
301  CGCCTCGTCT GA
```

This corresponds to the amino acid sequence <SEQ ID 104; ORF 018.a>:

a018.pep

```
  1  MQQGQLVGRV ARNKDMRNTG LHSQRIGNGY AARIFFDIDV FQTDIVNVRT
 51  AAYGCQHIFG NKYAFFAILL PMDFYIAVCV EFGLGFSIQM QFQFFTEHGF
101  RLV*
```

m018/a018 86.4% identity over a 103 aa overlap

```
                       10        20        30        40        50        60
m018.pep       MQQRQLVGRIACDEDMRNTGLHGQRVGNRYAARIFFDIDIFQTDIVNVRTAAHGCQHIFG
               ||| |||||:| ::||||||||:||:|| ||||||||||:||||||||||||:|||||||
a018           MQQGQLVGRVARNKDMRNTGLHSQRIGNGYAARIFFDIDVFQTDIVNVRTAAYGCQHIFG
                       10        20        30        40        50        60

                       70        80        90       100
m018.pep       NKYAFFAILLPMDFYIAVCIEFDLGFSIQMQFQFFAEHGVRLVX
               |||||||||||||||||||:|| ||||||||||||:||| ||||
a018           NKYAFFAILLPMDFYIAVCVEFGLGFSIQMQFQFFTEHGFRLVX
                       70        80        90       100
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 018 shows 84.5% identity over a 103 aa overlap with a predicted ORF (ORF 018.ng) from *N. gonorrhoeae:*

m018/g018

```
                       10        20        30        40        50        60
m018.pep       MQQRQLVGRIACDEDMRNTGLHGQRVGNRYAARIFFDIDIFQTDIVNVRTAAHGCQHIFG
               ||| |||||:| ::||||:||||||:|| ||||:| |||:|||||||||||::|||||||
g018           MQQGQLVGRVARNKDMRNAGLHGQRIGNGYAARV[illegible]DIDVFQTDIVNVRTATYGCQHIFG
                       10        20        30        40        50        60
```

```
                   70        80        90        100
m018.pep    NKYAFFAILLPMDFYIAVCIEFDLGFSIQMQFQFFAEHGVRLVX
            ||||||||||||||||||||:||||||||||||||||:||| ||||
g018        NKYAFFAILLPMDFYIAVCVEFDLGFSIQMQFQFFSEHGFRLVX
                   70        80        90        100
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 105>:

```
g019.seq (partial)
     1  ..ctgctggcgg ccctggtgct tgccgcgtgt tcttcgACAA ACAcacTGCC
    51    AGCCGGCAAG ACCCCGGCAG ACAATATAGA AActgcCgAC CTTTCGGCAA
   101    GCGTTCCCAC ccgcCCTGCC GAACCGGAAG GAAAAACGCT GGCAGATTAC
   151    GGCGGCTACC CGTCCGCACT GGATGCAGTG AAACAGAACA ACGATGCGGC
   201    AGCCGCCGCC TATTTGGAAA Acgcaggaga cagCGcgatg gcGGAAAatg
   251    tccgcaagga gtgGCTGa
```

This corresponds to the amino acid sequence <SEQ ID 106; ORF 019.ng>:

```
g019.pep (partial)
     1  ..LLAALVLAAC SSTNTLPAGK TPADNIETAD LSASVPTRPA EPEGKTLADY
    51    GGYPSALDAV KQNNDAAAAA YLENAGDSAM AENVRKEWL*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 107>:

```
m019.seq (partial)
     1  ATGTACCTAC CCTCTATGAA GCATTCCCTG CCGCTGCTGG CGGCCCTGGT
    51  GCTTGCCGCG TGTTCTTCGA CAAACACACT GCCAGCCGGC AAGACCCCGG
   101  CAGACAATAT AGAAACTGCC GACCTTTCGG CAAGCGTTCC CACCCGCCCT
   151  GCCGAACCCG AAAGAAAAAC GCTGGCAGAT TACGGCGGCT ACCCGTCCGC
   201  ACTGGATGCA GTGAAACAGA AAAACGATGC CGCCGTCGCC GCCTATTTGG
   251  AAAACGCCGG CGACAGCGCG ATGGCGGAAA ATGTCCGCAA CGAGTGGCTG
   301  AAGTCTTTGG GCGCACGCAG ACAGTGGACG CTGTTTGCAC AGGAATACGC
   351  CAAACTCGAA CCGGCAGGGC GCGCCCAAGA AGTCGAATGC TACGCCGATT
   401  CGAGCCGCAA CGACTATACG CGTGCCGCTG AACTGGTCAA AAATACGGGC
   451  AAACTGCCTT CGGGCTGCAC CAAACTGTTG GAACAGGCAG CCGCATCCGG
   501  CTTGTTGGAC GGCAACGACG CCTGGAGGCG CGTGCGCGGA CTGCTGGCCG
   551  GCCGCCAAAC CACAGACGCA CGCAACCTTG CCGCCGCATT GGGCAGCCCG
   601  TTTGACGGCG GTACACAAGG TTCGCGCGAA TATGCCCTGT TGAACGTCAT
   651  CGGCAAAGAA GCACGCAAAT CGCCGAATGC CGCCGCCCTG CTGTCCGAAA
   701  TGGAAAGCGG TTTAAGCCTC GAACAACGCA GTTTCGCGTG GGGCGTATTG
   751  GGGCATTATC AGTCGCAAAA CCTCAATGTG CCTGCCGCCT TGGACTATTA
   801  CGGCAAGGTT GCCGACCGCC GCCAACTGAC CGACGACCAA ATCGAGTGGT
   851  ACGCCCGCGC CGCCTTGCGC GCCCGACGTT GGGACGAGCT GGCCTCCGTT
   901  ATCTCGCATA TGCCCGAAAA ACTGCAAAAA AGCCCGACCT GGCTCTACTG
   951  GCTGGCACGC AGCCGCGCCG CAACGGGCAA CACGCAAGAG GCGGAAAAAC
  1001  TTTACAAACA GGCGGCAGCG ACGGGCAGGA ATTTTTATGC GGTGCTGGCA
  1051  GGGGAAGAAT TGGGTCGGAA AATCGATACG CGCAACAATG TGCCCGATGC
  1101  CGGCAAAAAC AGCGTCCGCC GCATGGCGGA AGACGGTGCA GTCAAACGCG
  1151  CACTGGTACT GTTCCAAAAC AGCCAATCTG CCGGTGATGC AAAAATGCGC
  1201  CGTCAGGCTC AGGCGGAATG GCGTTTTGCC ACACGCGGCT TTGACGAAGA
  1251  CAAGCTGCTG ACCGCCGCGC AAACCGCGTT CGACCACGGT TTTTACGATA
  1301  TGGCGGTCAA CAGCGCGGAA CGCACCGACC GCAAACTCAA CTACACCTTG
  1351  CGCTATATTT CGCCGTTTAA AGACACGGTA ATCCGCCACG CGCAAAATGT
  1401  TAATGTCGAT CCGGCTTGGG TTTATGGGCT GATTCGTCAG GAAAGCCGCT
  1451  TCGTTATAGG CGCGCAATCC CGCGTAGGCG CGCAGGGGCT GATGCAGGTT
  1501  ATGCCTGCCA CCGCGCGCGA AATCGCCGGC AAAATCGGTA TGGATGCCGC
  1551  ACAACTTTAC ACCGCCGACG GG...
```

This corresponds to the amino acid sequence <SEQ ID 108; ORF 019>:

```
m019.pep (partial)
     1  MYLPSMKHSL PLLAALVLAA CSSTNTLPAG KTPADNIETA DLSASVPTRP
```

```
 51  AEPERKTLAD YGGYPSALDA VKQKNDAAVA AYLENAGDSA MAENVRNEWL
101  KSLGARRQWT LFAQEYAKLE PAGRAQEVEC YADSSRNDYT RAAELVKNTG
151  KLPSGCTKLL EQAAASGLLD GNDAWRRVRG LLAGRQTTDA RNLAAALGSP
201  FDGGTQGSRE YALLNVIGKE ARKSPNAAAL LSEMESGLSL EQRSFAWGVL
251  GHYQSQNLNV PAALDYYGKV ADRRQLTDDQ IEWYARAALR ARRWDELASV
301  ISHMPEKLQK SPTWLYWLAR SRAATGNTQE AEKLYKQAAA TGRNFYAVLA
351  GEELGRKIDT RNNVPDAGKN SVRRMAEDGA VKRALVLFQN SQSAGDAKMR
401  RQAQAEWRFA TRGFDEDKLL TAAQTAFDHG FYDMAVNSAE RTDRKLNYTL
451  RYISPFKDTV IRHAQNVNVD PAWVYGLIRQ ESRFVIGAQS RVGAQGLMQV
501  MPATAREIAG KIGMDAAQLY TADG...
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 109>:

```
a019.seq
   1  ATGTACCCAC CCTCTCTGAA GCATTCCCTG CCGCTGCTGG TGGNCCTGGT
  51  GCTTGCCGCG TGTTCTTNGA CAAACACACT GTCAGCCGAC AAGACCCCGG
 101  CAGACAATAT AGAAACTGCC GACCTTTCGG CAAGCGTTCC CACCNGCCCT
 151  GCCGAACCCG AANGAAAAAC GTNGGCAGAT TACGGCGGCT ACCCGTCCGC
 201  ACTGGATGCA GTGAAACAGA AAAACGATGC CGCCGTCGCC GCCTATTTGG
 251  AAAACGCCGG CGACAGCGCG ATGGCGGAAA ATGTCCGCAA CGAGTGGCTG
 301  AAGTCTTTGG GCGCGCGCAG ACAGTGGACG CTGTNTGCAC ANGAATATGC
 351  NAAACTCGAA CCGGCANGGC GCGCCCAAGA AGTCGAATGC TACGCCGATT
 401  CGAGCCGCAA CGACTATACG CGTGCCGCCG AACTGGTCAA AAATACGGGC
 451  AAACTGCCTT CGGGCTGCAC CAAACTGTTG GAACAGGCAG CCGCATCCGG
 501  CTTGTTGGAC GGCAACGACG CCTGGAGGCG CGTGCGCGGA CTGCTGGCCG
 551  GCCGCCAAAC CACAGACGCA CGCAACCTTG CCGCCGCATT GGGCAGCCCG
 601  TTTGACGGCG GTACACAAGG TTCGCGCGAA TATGCCCTGT TGAACGTCAT
 651  CGGCAAAGAA GCACGCAAAT CGCCGAATGC CGCCGCCCTG CTGTCCGAAA
 701  TGGAAAGCGG TTTAAGCCTC GAACAACGCA GTTTCGCGTG GGCGTATTG
 751  GGGCATTATC AGTCGCAAAA CCTCAATGTG CCTGCCGCCT TGGACTATTA
 801  NGGCAAGGTT GCCGACCGCC GCCAACTGAC CGACGACCAA ATCGAGTGGT
 851  ACGCCCGCGC CGCNNTNNGC NNNCGNNGTT NGNANGANNT GGCNNCCGNN
 901  ANCNCGNNNN TGCNNGANAA ACNNNNNNAN AGNCNNANNT NGNTNNANTG
 951  NNTGGCACGC AGCCGCGCCG CNACGGGCAA CACGCAANAN GCGGANAAAC
1001  TNTACAAACA GGCGGCAGCA NCGGGCANGA ATTTTTATGC NGTGCTGNCN
1051  GGGGAAGAGT TGGGGCGCAN AATCGATACG CGCAACAATG TGCCCGATGC
1101  CGGCAAAANC AGCGTCCTCC GTATGGCGGA AGACGGCGCG ATTAAGCGCG
1151  CGCTGGTGCT GTTCCGAAAC AGCCGAACCG CCGGCGATGC GAAAATGCGC
1201  CGTCNGGCTC AGGCGGAATG GCGTTTCGCC ACACGCGGCT TCGATGAAGA
1251  CAAGCTGCTG ACCGCCGCGC AAACCGCGTT CGACCACGGT TTTTACGATA
1301  TGGCGGTCAA CAGCGCGGAA CGCACCGACC GCAAACTCAA CTACACCTTG
1351  CGCTACATTT CGNNNNNTNA NGACACGGTA ATCCGCCACG CGCAAAATGT
1401  TAATGTCGAT CCGGCGTGGG TTTACGGGCT GATTCGTCAG GAAAGCCGCT
1451  TCGTTATGGG CGCGCAATCC CGCGTAGGCG CGCAGGGGCT GATGCAGGTT
1501  ATGCCTGCCA CCGCGCGCGA AATCGCCGGC AAAATCGGTA TGGATGCCGC
1551  ACAACTTTAC ACCGCCGACG GCAATATCCG TATGGGGACG TGGTATATGG
1601  CGGACACCAA ACGCCGCCTG CAAAACAACG AAGTCCTCGC CACCGCAGGC
1651  TATAACGCCG GTCCCGGCAG GGCGCGCCGA TGGCAGGCGG ACACGCCCCT
1701  CGAAGGCGCG GTATATGCCG AAACCATCCC GTTTTCCGAA ACGCGCGACT
1751  ATGTCAAAAA AGTGATGGCC AATGCCGCCT ACTACGCCTC CCTCTTCGGC
1801  GCGCCGCACA TCCCGCTCAA ACAGCGTATG GGCATTGTCC CCGCCCGCTG
1851  A
```

This corresponds to the amino acid sequence <SEQ ID 110; ORF 019.a>:

```
a019.pep
  1  MYPPSLKHSL PLLVXLVLAA CSXTNTLSAD KTPADNIETA DLSASVPTXP
 51  AEPEXKTXAD YGGYPSALDA VKQKNDAAVA AYLENAGDSA MAENVRNEWL
101  KSLGARRQWT LXAXEYAKLE PAXRAQEVEC YADSSRNDYT RAAELVKNTG
151  KLPSGCTKLL EQAAASGLLD GNDAWRRVRG LLAGRQTTDA RNLAAALGSP
201  FDGGTQGSRE YALLNVIGKE ARKSPNAAAL LSEMESGLSL EQRSFAWGVL
251  GHYQSQNLNV PAALDYXGKV ADRRQLTDDQ IEWYARAAXX XRXXXXXAXX
301  XXXXXXKXXX XXXXXXXXAR SRAATGNTQX AXKLYKQAAA XGXNFYAVLX
351  GEELGRXIDT RNNVPDAGKX SVLRMAEDGA IKRALVLFRN SRTAGDAKMR
401  RXAQAEWRFA TRGFDEDKLL TAAQTAFDHG FYDMAVNSAE RTDRKLNYTL
```

```
451  RYISXXXDTV IRHAQNVNVD PAWVYGLIRQ ESRFVMGAQS RVGAQGLMQV
501  MPATAREIAG KIGMDAAQLY TADGNIRMGT WYMADTKRRL QNNEVLATAG
551  YNAGPGRARR WQADTPLEGA VYAETIPFSE TRDYVKKVMA NAAYYASLFG
601  APHIPLKQRM GIVPAR*
```

m019/a019 88.9% identity over a 524 aa overlap

```
                    10        20        30        40        50        60
m019.pep    MYLPSMKHSLPLLAALVLAACSSTNTLPAGKTPADNIETADLSASVPTRPAEPERKTLAD
            || ||:|||||||: ||||||| |||| | |||||||||||||||||| ||||| || ||
a019        MYPPSLKHSLPLLVXLVLAACSXTNTLSADKTPADNIETADLSASVPTXPAEPEXKTXAD
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m019.pep    YGGYPSALDAVKQKNDAAVAAYLENAGDSAMAENVRNEWLKSLGARRQWTLFAQEYAKLE
            ||||||||||||||||||||||||||||||||||||||||||||||||||| | ||||||
a019        YGGYPSALDAVKQKNDAAVAAYLENAGDSAMAENVRNEWLKSLGARRQWTLXAXEYAKLE
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m019.pep    PAGRAQEVECYADSSRNDYTRAAELVKNTGKLPSGCTKLLEQAAASGLLDGNDAWRRVRG
            || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a019        PAXRAQEVECYADSSRNDYTRAAELVKNTGKLPSGCTKLLEQAAASGLLDGNDAWRRVRG
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m019.pep    LLAGRQTTDARNLAAALGSPFDGGTQGSREYALLNVIGKEARKSPNAAALLSEMESGLSL
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a019        LLAGRQTTDARNLAAALGSPFDGGTQGSREYALLNVIGKEARKSPNAAALLSEMESGLSL
                   190       200       210       220       230       240

                   250       260       270       280       290       300
m019.pep    EQRSFAWGVLGHYQSQNLNVPAALDYYGKVADRRQLTDDQIEWYARAALRARRWDELASV
            |||||||||||||||||||||||||| |||||||||||||||||||||   |     |  
a019        EQRSFAWGVLGHYQSQNLNVPAALDYXGKVADRRQLTDDQIEWYARAAXXXRXXXXXAXX
                   250       260       270       280       290       300

                   310       320       330       340       350       360
m019.pep    ISHMPEKLQKSPTWLYWLARSRAATGNTQEAEKLYKQAAATGRNFYAVLAGEELGRKIDT
                  |     :     ||||||||||| | ||||||||:| |||||| |||||| |||
a019        XXXXXXKXXXXXXXXXXXXARSRAATGNTQXAXKLYKQAAAXGXNFYAVLXGEELGRXIDT
                   310       320       330       340       350       360

                   370       380       390       400       410       420
m019.pep    RNNVPDAGKNSVRRMAEDGAVKRALVLFQNSQSAGDAKMRRQAQAEWRFATRGFDEDKLL
            ||||||||| || |||||||:|||||||:||::|||||||| ||||||||||||||||||
a019        RNNVPDAGKXSVLRMAEDGAIKRALVLFRNSRTAGDAKMRRXAQAEWRFATRGFDEDKLL
                   370       380       390       400       410       420

                   430       440       450       460       470       480
m019.pep    TAAQTAFDHGFYDMAVNSAERTDRKLNYTLRYISPFKDTVIRHAQNVNVDPAWVYGLIRQ
            ||||||||||||||||||||||||||||||||||   |||||||||||||||||||||||
a019        TAAQTAFDHGFYDMAVNSAERTDRKLNYTLRYISXXXDTVIRHAQNVNVDPAWVYGLIRQ
                   430       440       450       460       470       480

                   490       500       510       520
m019.pep    ESRFVIGAQSRVGAQGLMQVMPATAREIAGKIGMDAAQLYTADG
            |||||:||||||||||||||||||||||||||||||||||||||
a019        ESRFVMGAQSRVGAQGLMQVMPATAREIAGKIGMDAAQLYTADGNIRMGTWYMADTKRRL
                   490       500       510       520       530       540

a019        QNNEVLATAGYNAGPGRARRWQADTPLEGAVYAETIPFSETRDYVKKVMANAAYYASLFG
                   550       560       570       580       590       600
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 019 shows 95.5% identity over a 89 aa overlap with a predicted ORF (ORF 019.ng) from *N. gonorrhoeae:*

```
g019/m019

                              10        20        30        40       49
g019.pep               LLAALVLAACSSTNTLPAGKTPADNIETADLSASVPTRPAEPEGKTLAD
                       ||||||||||||||||||||||||||||||||||||||||||| |||||
m019        MYLPSMKHSLPLLAALVLAACSSTNTLPAGKTPADNIETADLSASVPTRPAEPERKTLAD
                    10        20        30        40        50        60

            50        60        70        80       89
g019.pep    YGGYPSALDAVKQNNDAAAAAYLENAGDSAMAENVRKEWL
            |||||||||||||:||||:|||||||||||||||||:|||
m019        YGGYPSALDAVKQKNDAAVAAYI ENAGDSAMAENVRNEWLKSLGARRQWTLFAQEYAKLE
                    70        80        90       100       110       120
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 111>:

```
g023.seq
      1  ATGGTAGAAC GTAAATTGAC CGGTGCCCAT TACGGTTTGC GCGATTGGGT
     51  AATGCAGCGT GCGACTGCGG TTATTATGTT GATTTATACC GTTGCACTTT
    101  TAGTGGTTCT ATTTGCCCTG CCTAAAGAAT ATCCGGCATG GCAGGCATTT
    151  TTTAGTCAAG CTTGGGTAAA AGTATTTACC CAAGTGAGCT TTATCGCCGT
    201  ATTCTTGCAC GCTTGGGTGG GTATCCGCGA TTTGTGGATG GACTATATCA
    251  AACCCTTCGG CGTGCGTTTG TTTTTGCAGG TTGCCACCAT TGtctGGCTG
    301  GTCGGCTGCC TCGTGTATTC AGTTAAAGTG ATTTGGGGGT AA
```

This corresponds to the amino acid sequence <SEQ ID 112; ORF 023.ng>:

```
g023.pep
      1  MVERKLTGAH YGLRDWVMQR ATAVIMLIYT VALLVVLFAL PKEYPAWQAF
     51  FSQAWVKVFT QVSFIAVFLH AWVGIRDLWM DYIKPFGVRL FLQVATIVWL
    101  VGCLVYSVKV IWG*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 113>:

```
m023.seq
      1  ATGGTAGAAC GTAAATTGAC CGGTGCCCAT TACGGTTTGC GCGATTGGGT
     51  GATGCAACGT GCGACTGCGG TTATTATGTT GATTTATACC GTTGCACTTT
    101  TAGTGGTTCT ATTTTCCCTG CCTAAAGAAT ATTCGGCATG GCAGGCATTT
    151  TTTAGTCAAA CTTGGGTAAA AGTATTTACC CAAGTGAGCT TCATCGCCGT
    201  ATTCTTGCAC GCTTGGGTGG GTATCCGCGA TTTGTGGATG GACTATATCA
    251  AACCCTTCGG CGTGCGTTTG TTTTTGCAGG TTGCCACCAT CGTTTGGCTG
    301  GTCGGCTGTC TCGTGTATTC AGTTAAAGTG ATTTGGGGGT AA
```

This corresponds to the amino acid sequence <SEQ ID 114; ORF 023>:

```
m023.pep
      1  MVERKLTGAH YGLRDWVMQR ATAVIMLIYT VALLVVLFSL PKEYSAWQAF
     51  FSQTWVKVFT QVSFIAVFLH AWVGIRDLWM DYIKPFGVRL FLQVATIVWL
    101  VGCLVYSVKV IWG*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 115>:

```
a023.seq
      1  ATGGTAGAAC GTAAATTGAC CGGTGCCCAT TACGGTTTGC GGGATTGGGC
     51  GATGCAACGT GCGACCGCGG TTATTATGTT GATTTATACC GTTGCACTTT
    101  TAGTGGTTCT ATTTGCTCTG CCTAAAGAAT ATTCGGCATG GCAGGCATTT
    151  TTTAGTCAAA CTTGGGTAAA AGTATTTACC CAAGTGAGCT TCATCGCCGT
```

```
201  ATTCTTGCAC GCTTGGGTGG GTATCCGCGA TTTGTGGATG GACTATATNA
251  AACCCTTCGG CGTGCGTTTG TTTTTGCAGG TTGCCACCAT CGTCTGGCTG
301  GTCGGCTGCT TGGTGTATTC AATTAAAGTA ATTTGGGGGT AA
```

This corresponds to the amino acid sequence <SEQ ID 116; ORF 023.a>:

```
a023.pep
  1  MVERKLTGAH YGLRDWAMQR ATAVIMLIYT VALLVVLFAL PKEYSAWQAF
 51  FSQTWVKVFT QVSFIAVFLH AWVGIRDLWM DYXKPFGVRL FLQVATIVWL
101  VGCLVYSIKV IWG*
```

m023/a023 96.5% identity over a 113 aa overlap

```
                   10        20        30        40        50        60
m023.pep   MVERKLTGAHYGLRDWVMQRATAVIMLIYTVALLVVLFSLPKEYSAWQAFFSQTWVKVFT
           ||||||||||||||||:||||||||||||||||||||||:||||||||||||||||||||
a023       MVERKLTGAHYGLRDWAMQRATAVIMLIYTVALLVVLFALPKEYSAWQAFFSQTWVKVFT
                   10        20        30        40        50        60

                   70        80        90       100       110
m023.pep   QVSFIAVFLHAWVGIRDLWMDYIKPFGVRLFLQVATIVWLVGCLVYSVKVIWGX
           |||||||||||||||||||||| ||||||||||||||||||||||||:||||||
a023       QVSFIAVFLHAWVGIRDLWMDYXKPFGVRLFLQVATIVWLVGCLVYSIKVIWGX
                   70        80        90       100       110
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 023 shows 97.3% identity over a 113 aa overlap with a predicted ORF (ORF 023.ng) from *N. gonorrhoeae:*

```
g023/m023

                   10        20        30        40        50        60
g023.pep   MVERKLTGAHYGLRDWVMQRATAVIMLIYTVALLVVLFALPKEYPAWQAFFSQAWVKVFT
           |||||||||||||||||||||||||||||||||||||||:||||| ||||||||:|||||
m023       MVERKLTGAHYGLRDWVMQRATAVIMLIYTVALLVVLFSLPKEYSAWQAFFSQTWVKVFT
                   10        20        30        40        50        60

                   70        80        90       100       110
g023.pep   QVSFIAVFLHAWVGIRDLWMDYIKPFGVRLFLQVATIVWLVGCLVYSVKVIWGX
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
m023       QVSFIAVFLHAWVGIRDLWMDYIKPFGVRLFLQVATIVWLVGCLVYSVKVIWGX
                   70        80        90       100       110
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 117>:

```
g025.seq
  1  ATGTTGAAAC AAAcgACACT TTTGGCAGCT TGTACCGCCG TTGCCGCTCT
 51  GTTGGGCGGT TGcgCCACCC AACAGCCTGC TccTGTCATT GCAGGCAATT
101  CAGGTATGCA GACCGTATCG TCTGCGCCGG TTTACAATCC TTATGGCGCA
151  ACGCCGTACA ATGCCGCTCC TGCCGCCAac gatgcGCCgT ATGTGCCGCC
201  CGTGCAAact gcgccggttT ATTCGCCTCC TGCTTATGTT CCGCcgtCTG
251  CACCTGCCGT TTCGGgtaca tatgtTCCTT CTTACGCACC CgtcgACATC
301  aacgCGGCGa cgCataCTAT TGTGCGTGGC GACACgGtgt acaACATTTc
351  caaAcgCtac CATATCTCTC AAGACGATTT CCGTGCGTGG AACGGCATGA
401  CCGACAATAC GTTGAGCATC CGTCAGATTG TTAAAGTCAA ACCGGCaggA
451  TATGCCGCAC CGAAAACCGC AGCCGTAGAA AGCAGGCCCG CCGTACCGGC
501  TGCCGCGCAA ACCCCTGTGA AACCCGCCGC gcaACCGCCC GTTCAGTCCG
551  CGCCGCAACC TGCCGCGCCC GCTGCGGAAA ATAAAGCGGT TCCCGCCCCC
601  GCGCCCGCCC CGCAATCTCC TGCCGCTTCG CCTTCCGGCA CGCGTTCGGT
651  CGGCGGCATT GTTTGGCAGC GTCCGACCCA AGGTAAAGTG GTTGCCGATT
```

```
 701  TCGGCGGCGG CAACAAGGGT GTCGATATTG CCGGCAATGC CGGACAACCC
 751  GTTTTGGCGG CGGCTGACGG CAAAGTGGTT TATGCCGGTT CAGGTTTGAG
 801  GGGATACGGA AACTTGGTCA TCATCCAGCA CAATTCCTCT TTCCTGACCG
 851  CGTACGGGCA CAACCAAAAA TTGCTGGTCG GCGAAGGTCA GCAGGTCAAA
 901  CGCGGTCAGC AGGTTGCTTT GATGGGTAAT ACCGATGCTT CCAGAACGCA
 951  GCTTCATTTC GAGGTGCGTC AAAACGGCAA ACCGGTTAAC CCGAACAGCT
1001  ATATCGCGTT CTGA
```

This corresponds to the amino acid sequence <SEQ ID 118; ORF 025.ng>:

```
g025.pep
   1  MLKQTTLLAA CTAVAALLGG CATQQPAPVI AGNSGMQTVS SAPVYNPYGA
  51  TPYNAAPAAN DAPYVPPVQT APVYSPPAYV PPSAPAVSGT YVPSYAPVDI
 101  NAATHTIVRG DTVYNISKRY HISQDDFRAW NGMTDNTLSI GQIVKVKPAG
 151  YAAPKTAAVE SRPAVPAAAQ TPVKPAAQPP VQSAPQPAAP AAENKAVPAP
 201  APAPQSPAAS PSGTRSVGGI VWQRPTQGKV VADFGGGNKG VDIAGNAGQP
 251  VLAAADGKVV YAGSGLRGYG NLVIIQHNSS FLTAYGHNQK LLVGEGQQVK
 301  RGQQVALMGN TDASRTQLHF EVRQNGKPVN PNSYIAF*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 119>:

```
m025.seq  (partial)
   1  ..GTGCCGCCGG TGCAAAGCGC GCCGGTTTAT ACGCCTCCTG CTTATGTTCC
  51    GCCGTCTGCA CCTGCCGTTT CGGGTACATA CGTTCCTTCT TACGCACCCG
 101    TCGACATCAA CGCGGCGACG CATACTATTG TGCGCGGCGA CACGGTGTAC
 151    AACATTTCCA AACGCTACCA TATCTCTCAA GACGATTTCC GTGCGTGGAA
 201    CGGCATGACC GACAATACGT TGAGCATCGG TCAGATTGTT AAAGTCAAAC
 251    CGGCAGGATA TGCCGCACCG AAAGCCGCAG CCGTAAAAAG CAGGCCCGCC
 301    GTACCGGCTG CCGCGCAACC GCCCGTACAG TCCGCACCCG TCGACATTAA
 351    CGCGGCGACG CATACTATTG TGCGCGGCGA CACGGTGTAC AACATTTCCA
 401    AACGCTACCA TATCTCTCAA GACGATTTCC GTGCGTGGAA CGGCATGACC
 451    GACAATATGT TGAGCATCGG TCAGATTGTT AAAGTCAAAC CGGCAGGATA
 501    TGCCGCACCG AAAACCGCAG CCGTAGAAAG CAGGCCCGCC GTACCGGCTG
 551    CCGTGCAAAC CCCTGTGAAA CCCGCCGCGC AACCGCCTGT GCAGTCCGCG
 601    CCGCAACCTG CCGCGCCCGC TGCGGAAAAT AAAGCGGTTC CCGCGCCCGC
 651    CCCGCAATCT CCTGCCGCTT CGCCTTCCGG CACGCGTTCG GTCGGCGGCA
 701    TTGTTTGGCA GCGTCCGACG CAAGGTAAAG TGGTTGCCGA TTTCGGCGGC
 751    AACAACAAGG GTGTCGATAT TGCCGGTAAT GCGGGACAGC CCGTTTTGGC
 801    GGCGGCTGAC GGCAAAGTGG TTTATGCCGG TTCAGGTTTG AGGGGATACG
 851    GAAACTTGGT CATCATCCAG CATAATTCTT CTTTCCTGAC CGCATACGGG
 901    CACAACCAAA AATTGCTGGT CGGCGAGGGG CAGCAGGTCA AACGCGGTCA
 951    GCAGGTTGCT TTGATGGGCA ATACCGATGC TTCCAGAACG CAGCTTCATT
1001    TCGAGGTGCG TCAAAACGGC AAACCGGTTA ACCCGAACAG CTATATCGCG
1051    TTCTGA
```

This corresponds to the amino acid sequence <SEQ ID 110; ORF 025>:

```
m025.pep  (partial)
   1  ..VPPVQSAPVY TPPAYVPPSA PAVSGTYVPS YAPVDINAAT HTIVRGDTVY
  51    NISKRYHISQ DDFRAWNGMT DNTLSIGQIV KVKPAGYAAP KAAAVKSRPA
 101    VPAAAQPPVQ SAPVDINAAT HTIVRGDTVY NISKRYHISQ DDFRAWNGMT
 151    DNMLSIGQIV KVKPAGYAAP KTAAVESRPA VPAAVQTPVK PAAQPPVQSA
 201    PQPAAPAAEN KAVPAPAPQS PAASPSGTRS VGGIVWQRPT QGKVVADFGG
 251    NNKGVDIAGN AGQPVLAAAD GKVVYAGSGL RGYGNLVIIQ HNSSFLTAYG
 301    HNQKLLVGEG QQVKRGQQVA LMGNTDASRT QLHFEVRQNG KPVNPNSYIA
 351    F*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 111>:

```
a025.seq
   1  ATGTTGACAC CAACAACACT TTAGGTAGCT TGTACCGCCC TTGCCGCTCA
  51  GTTGGGCGGA TGCCCCACCC AACACCCTTC TCCTGTCATT GCAGGCAATT
 101  CAGGTATGCA GACCGTACCG TCTGCGCCGG TTTACAATCC TTATGGCGCA
```

```
 151  ACGCCGTACA ATGCCGCTCC TGCCGCCAAC GATGCGCCGT ATGTGCCGCC
 201  GGTGCAAAGC GCGCCGGTTT ATANGCCTCC TGCTTATGTT CCGCCGTCTG
 251  CACCTGCCGT TTCGGGTACA TACGTTCCTT CTTACGCANC CGTCGACATC
 301  AACGCGGCGA CCCATACTAT TGTGCGCGGC GACACCGTGT ACAAGATTTC
 351  CAAATGCTAC CATATCTCTC AAGACGATTT CCGTGCGTGG AACGGCATGA
 401  CCGACAATAC GTTGAGCATC GGTCAGATTG TTAAAGTCAA ACCGGCAGGA
 451  TATGCCGCAC CGAAAGCCGC AGCCGTAAAA AGCAGGCCCG CCGTACCGGC
 501  TGCCGCGCAA CCGCTCGTAC AGTCCGCACC CGTCGACATC AACGCGGCGA
 551  CGCATACTAT TGTGCGCGGC GACACGGTGT ACAACATTTC CAAACGCTAC
 601  CATATCTCTC AAGACGATTT CCGTGCGTGG AACGGCATGA CCGACAATAC
 651  GTTGAGCATC GGTCAGATTG TTAAAGTCAA ACCGGCAGGA TATGCCGCAC
 701  CGAAAGCCGC AGCCGTAAAA AGCAGGCCCG CCGTACCGGC TGCCGTGCAA
 751  ACCCCTGTGA AACCCGCCGC GCAACCGCCT GTGCAGTCCG CGCCGCAACC
 801  TGCCGCGCCC GCTGCGGAAA ATAAAGCGGT TCCCGCGCCC GCCCCGCAAT
 851  CTCCTGCCGC TTCGCCTTCC GGCACGCGTT CGGTCGGCGG CATTGTTTGG
 901  CAGCGTCCGA CGCAAGGTAA AGTGGTTGCC GATTTCGGCG GCAACAACAA
 951  GGGTGTCGAT ATTGCAGGAA ATGCGGGACA GCCCGTTTTG GCGGCGGCTG
1001  ACGGCAAAGT GGTTTATGCA GGTTCCGGTT TGAGGGGATA CGGCAATTTG
1051  GTCATCATCC AGCATAATTC TTCCTTCCTG ACCGCATACG GGCACAACCA
1101  AAAATTGCTG GTCGGCGAAG GCCAGCAGGT CAAACGCGGG CAGCAGGTCG
1151  CTTTGATGGG CAATACCGAG GCTTCTAGAA CGCAGCTTCA TTTCGAGGTG
1201  CGGCAAAACG GCAAACCGGT TAATCCGAAC AGCTATATCG CGTTCTGA
```

This corresponds to the amino acid sequence <SEQ ID 112; ORF 025.a>:

```
a025.pep
   1  MLTPTTL*VA CTALAAQLGG CPTQHPSPVI AGNSGMQTVP SAPVYNPYGA
  51  TPYNAAPAAN DAPYVPPVQS APVYXPPAYV PPSAPAVSGT YVPSYAXVDI
 101  NAATHTIVRG DTVYKISKCY HISQDDFRAW NGMTDNTLSI GQIVKVKPAG
 151  YAAPKAAAVK SRPAVPAAAQ PLVQSAPVDI NAATHTIVRG DTVYNISKRY
 201  HISQDDFRAW NGMTDNTLSI GQIVKVKPAG YAAPKAAAVK SRPAVPAAVQ
 251  TPVKPAAQPP VQSAPQPAAP AAENKAVPAP APQSPAASPS GTRSVGGIVW
 301  QRPTQGKVVA DFGGNNKGVD IAGNAGQPVL AAADGKVVYA GSGLRGYGNL
 351  VIIQHNSSFL TAYGHNQKLL VGEGQQVKRG QQVALMGNTE ASRTQLHFEV
 401  RQNGKPVNPN SYIAF*
```

m025/a025 97.4% identity over a 351 aa overlap

```
                                                    10        20        30
m025.pep                                  VPPVQSAPVYTPPAYVPPSAPAVSGTYVPS
                                          ||||||||||:|||||||||||||||||||
a025        GMQTVPSAPVYNPYGATPYNAAPAANDAPYVPPVQSAPVYXPPAYVPPSAPAVSGTYVPS
                    40        50        60        70        80        90

                    40        50        60        70        80        90
m025.pep    YAPVDINAATHTIVRGDTVYNISKRYHISQDDFRAWNGMTDNTLSIGQIVKVKPAGYAAP
            || |||||||||||||||||:||| |||||||||||||||||||||||||||||||||||
a025        YAXVDINAATHTIVRGDTVYKISKCYHISQDDFRAWNGMTDNTLSIGQIVKVKPAGYAAP
                   100       110       120       130       140       150

                   100       110       120       130       140       150
m025.pep    KAAAVKSRPAVPAAAQPPVQSAPVDINAATHTIVRGDTVYNISKRYHISQDDFRAWNGMT
            |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
a025        KAAAVKSRPAVPAAAQPLVQSAPVDINAATHTIVRGDTVYNISKRYHISQDDFRAWNGMT
                   160       170       180       190       200       210

                   160       170       180       190       200       210
m025.pep    DNMLSIGQIVKVKPAGYAAPKTAAVESRPAVPAAVQTPVKPAAQPPVQSAPQPAAPAAEN
            || |||||||||||||||||:|||:|||||||||||||||||||||||||||||||||||
a025        DNTLSIGQIVKVKPAGYAAPKAAAVKSRPAVPAAVQTPVKPAAQPPVQSAPQPAAPAAEN
                   220       230       240       250       260       270

                   220       230       240       250       260       270
m025.pep    KAVPAPAPQSPAASPSGTRSVGGIVWQRPTQGKVVADFGGNNKGVDIAGNAGQPVLAAAD
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
a025        KAVPAPAPQSPAASPSGTRSVGGIVWQRPTQGKVVADFGGNNKGVDIAGNAGQPVLAAAD
               280       290       300       310       320       330

                      280       290       300       310       320       330
m025.pep    GKVVYAGSGLRGYGNLVIIQHNSSFLTAYGHNQKLLVGEGQQVKRGQQVALMGNTDASRT
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
a025        GKVVYAGSGLRGYGNLVIIQHNSSFLTAYGHNQKLLVGEGQQVKRGQQVALMGNTEASRT
               340       350       360       370       380       390

                      340       350
m025.pep    QLHFEVRQNGKPVNPNSYIAFX
            ||||||||||||||||||||||
a025        QLHFEVRQNGKPVNPNSYIAFX
               400       410
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 025 shows 75.6% identity over a 353 aa overlap with a predicted ORF (ORF 025.ng) from *N. gonorrhoeae:*

```
m025/g025

                                                    10        20        30
m025.pep                                  VPPVQSAPVYTPPAYVPPSAPAVSGTYVPS
                                          |||||:|||||:||||||||||||||||||
g025        GMQTVSSAPVYNPYGATPYNAAPAANDAPYVPPVQTAPVYSPPAYVPPSAPAVSGTYVPS
               40        50        60        70        80        90

                      40        50        60        70        80        90
m025.pep    YAPVDINAATHTIVRGDTVYNISKRYHISQDDFRAWNGMTDNTLSIGQIVKVKPAGYAAP
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g025        YAPVDINAATHTIVRGDTVYNISKRYHISQDDFRAWNGMTDNTLSIGQIVKVKPAGYAAP
               100       110       120       130       140       150

                     100       110       120       130       140       150
m025.pep    KAAAVKSRPAVPAAAQPPVQSAPVDINAATHTIVRGDTVYNISKRYHISQDDFRAWNGMT
            |
g025        K-----------------------------------------------------------

                     160       170       180       190       200       210
m025.pep    DNMLSIGQIVKVKPAGYAAPKTAAVESRPAVPAAVQTPVKPAAQPPVQSAPQPAAPAAEN
                                ||||||||||||:|||||||||||||||||||||||||
g025        --------------------TAAVESRPAVPAAAQTPVKPAAQPPVQSAPQPAAPAAEN
                                  160       170       180       190

                      220       230       240       250       260
m025.pep    KAVPAPAP--QSPAASPSGTRSVGGIVWQRPTQGKVVADFGGNNKGVDIAGNAGQPVLAA
            ||||||||  ||||||||||||||||||||||||||||||||:|||||||||||||||||
g025        KAVPAPAPAPQSPAASPSGTRSVGGIVWQRPTQGKVVADFGGGNKGVDIAGNAGQPVLAA
               200       210       220       230       240       250

           270       280       290       300       310       320
m025.pep    ADGKVVYAGSGLRGYGNLVIIQHNSSFLTAYGHNQKLLVGEGQQVKRGQQVALMGNTDAS
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g025        ADGKVVYAGSGLRGYGNLVIIQHNSSFLTAYGHNQKLLVGEGQQVKRGQQVALMGNTDAS
               260       270       280       290       300       310

           330       340       350
m025.pep    RTQLHFEVRQNGKPVNPNSYIAFX
```

```
              ||||||||||||||||||||||||
g025          RTQLHFEVRQNGKPVNPNSYIAFX
                 320       330
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 113>:

g031.seq

```
  1  ATGGTGTCCC TCCGCTTCAG ATTCGGCAAC CACTTTAAAC GCCGACATTC
 51  TGACAATTTC CTTTTCCGCC AGCCAAATAT CATGCGTATC TTTCGGTTCG
101  GGCTTGTTGG GCATGGCAAC CTTCAACAGC CGCGCCATCA CAGGAATCGT
151  CGTTCCCTGA ATCAGCAGCG ACAGCACCAC CACGGCAAAC GCCACATCAA
201  ACAGCAGGTG CGAATTGGGA ACGCCCATCA CCAGCGGCAT CATCGCCAGC
251  GAAATCGGTA CGGCTCCTCG CAAGCCCAAC CAACTGATAT ACGCCTTTTC
301  ACGCAGGCTG TAATTGAATT TCCACAAACC GCCGAACACT GCCAGCGGAC
351  GCGCGACCAG CATCAGGAAC GCCGCAATCG CCAAGGCTTC CGCCGCCCTG
401  TCCAACACGC CGGCGGGAGA AACCAGCAGA CCGAGCATGA CGAACAAAGT
451  TGCCTGCGCC AGCCAAGCCA AACCGTCCAT CACACGCAAA ACGTGTTCCG
501  TcgcACGGTT GCGCTGGTTA CCGACAATGA TGCCGGCAAG GTAAACCGCC
551  AAAAAGCCGC TGCCGCCTAT GGTATTGGTA AACGCAAACA CAAGCAGCCC
601  GCCCGACACA ATCATCAGCG CGTACAGACC TTCCGtacac acctccaatt
651  cccaatcaac gtcatagctg tctcccgtgt taaaatgttc ttcacttcag
701  aatccccccc ttcttcccag cccgaaacct tcatgtgtta naccctgggg
751  tgccccaacg gatttagtaa cctcccaatg actctgcttg tcgcccccctt
801  cgcccgcttt ctccttccgg gaaaacttgt tgtccccgtc ttacattaa
```

This corresponds to the amino acid sequence <SEQ ID 114; ORF 031.ng>:

g031.pep

```
  1  MVSLRFRFGN HFKRRHSDNF LFRQPNIMRI FRFGLVGHGN LQQPRHHRNR
 51  RSLNQQRQHH HGKRHIKQQV RIGNAHHQRH HRQRNRYGSS QAQPTDIRLF
101  TQAVIEFPQT AEHCQRTRDQ HQERRNRQGF RRPVQHAGGR NQQTEHDEQS
151  CLRQPSQTVH HTQNVFRRTV ALVTDNDAGK VNRQKAAAAY GIGKRKHKQP
201  ARHNHQRVQT FRTHLQFPIN VIAVSRVKMF FTSESPPSSQ PETFMCXTLG
251  CPNGFSNLPM TLLVAPFARF LLPGKLVVPV LH*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 115>:

m031.seq (partial)

```
  1  ...CGCCTGAAGC ACGGTGTCGG ACTGCATTTC TATTCGGCTA TACGCCTTTT
 51     CACGCAGGCT GTAATTGAAT TTCCACAAAC CGCCGAACAC TGCCGACGGA
101     CGCGCGACCA GCATCAGGAA CGCCGCAATC GCCAAgGCTT CCGCCGCCCT
151     GTCCAACACG TTGGCAGGAG AAACCAGCAG CAAAGGCATT CCCAAACGTG
201     CGGACAAAGT GGTCGAAACC ACGCTCAGAA ACAACAGTGC GCCACCCGGC
251     AG....
```

This corresponds to the amino acid sequence <SEQ ID 116; ORF 031>:

m031.pep (partial)

```
  1  ...RLKHGVGLHF YSAIRLFTQA VIEFPQTAEH CRRTRDQHQE RRNRQGFRRP
 51     VQHVGRRNQQ QRHSQTCGQS GRNHAQKQQC ATRQ....
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 117>:

a031.seq

```
  1  ATACGCCTTT TCACGCAGGC TGTAATTGAA TTTCCACAAA CCGCCGAACA
 51  CTGCCGGCGG ACGCGCGACC AGCATCAGGA ACGCCGCAAT CGCCAAGGCT
101  TCCGCCGCCC CGTCCAACAC GTTGGCAGGA GAAACCAGCA GCAAAGGCAT
151  TCCCAAACGT GCGGACAAAG TGGTCGAAAC CACGCTCAGA AACAACAGTG
201  CGCCACCCGG CAG
```

This corresponds to the amino acid sequence <SEQ ID 118; ORF 031.a>:

```
a031.pep  (partial)
  1  IRLFTQAVIE FPQTAEHCRR TRDQHQERRN RQGFRRPVQH VGRRNQQQRH
 51  SQTCGQSGRN HAQKQQCATR Q
```

m031/a031 100.0% identity over a 71 aa overlap

```
                   10        20        30        40        50        60
m031.pep   RLKHGVGLHFYSAIRLFTQAVIEFPQTAEHCRRTRDQHQERRNRQGFRRPVQHVGRRNQQ
                        ||||||||||||||||||||||||||||||||||||||||||||||
a031                    IRLFTQAVIEFPQTAEHCRRTRDQHQERRNRQGFRRPVQHVGRRNQQ
                                10        20        30        40

                   70        80
m031.pep   QRHSQTCGQSGRNHAQKQQCATRQ
           ||||||||||||||||||||||||
a031       QRHSQTCGQSGRNHAQKQQCATRQ
            50        60        70
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

ORF 031 shows 60.0% identity over a 85 aa overlap with a predicted ORF (ORF 031.ng) from *N. gonorrhoeae:*

m031/g031

```
                                                  10        20        30
m031.pep                                RLKHGVGLHFYSAIRLFTQAVIEFPQTAEH
                                        | ::| :     : ||:|||||||||||||
g031       NQQRQHHHGKRHIKQQVRIGNAHHQRHHRQRNRYGSSQAQPTDIRLFTQAVIEFPQTAEH
               60        70        80        90       100       110

                   40        50        60        70        80
m031.pep   CRRTRDQHQERRNRQGFRRPVQHVGRRNQQQRHS-QTCGQSGRNHAQKQQCATRQ
           |:||||||||||||||||||||||:| |||| :|: |:| ::   : ::: | : |:
g031       CQRTRDQHQERRNRQGFRRPVQHAGGRNQQTEHDEQSCLRQPSQTVHHTQNVFRRTVALV
              120       130       140       150       160       170

g031       TDNDAGKVNRQKAAAAYGIGKRKHKQPARHNHQRVQTFRTHLQFPINVIAVSRVKMFFTS
              180       190       200       210       220       230
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 119>:

g032.seq

```
  1 ATGCGGCGAA ACGTGCCTGC CGTCGCCGTA TTGCGCCGCC CACGATTCGA
 51 GGCGTTTTTG GATTTGGCGT TGGCTCAGGC GCGTGCCGTT CCTGCCGGTA
101 AACAGGGCTT TGCCGTCCGA TGCCGTCTGA CGCAGCGGCA GATAGTTTTT
151 CAAGGCTTCC ACGCTTTTGC CGGTCAGCGG AACCTGACGC TGCTTGCGCC
201 CTTTGCCGGT AACGTGTACC CACGCTTCGT CCAAATATAC ATCATCTGCA
251 TTCAAGCCGT GTATCTCGCT CACGCGCAAA CCGCTGCCGT ACATCAGCTC
301 GAACAGCGCG TGGTCGCGCA CCGCCAGCGG GTCGCCGCCG TCCACGGGCA
351 AATCCAACAT CCGGTTCAGC CATTCCTGCG GCAGGGCTTT GGGTACGCGC
401 TCGGGCTGCT TCGGCGGTTT GATGTCGGCG GTCGGGTCGG CGCGCATCAG
451 CCCGCGTTTG ACCAGCCAGG CGCAATACTG CCGCCACGCC GACAGCTTGC
501 GCGCCAGCGT CCGACCGTCC AAACCGCGCT GCGACAGCCG CCGCAACGCC
551 GccgTAAAAT CGCGCCGCGA CAAGTCCTGC GGCACGCCgc ctgcaTCTTC
601 AGACGGCATT TGTGCCAACA GTGCAAACAG TTCTTCCAAA TCGCGCCGGT
651 ATGCCGCAAC CGTGTGCTCC GACTTGCCCT CGCGCACGAT GTTTTCCAAA
701 TAAGCGTCAA AATacgccgC AAACccgTCC AAAACCATAA CCGTCCCACA
751 CAAATATCAA AAAACCAGTG A
```

This corresponds to the amino acid sequence <SEQ ID 110; ORF 032.ng>:

g032.pep

```
  1 MRRNVPAVAV LRRPRFEAFL DLALAQARAV PAGKQGFAVR CRLTQRQIVF
 51 QGFHAFAGQR NLTLLAPFAG NVYPRFVQIY IICIQAVYLA HAQTAAVHQL
```

```
101  EQRVVAHRQR VAAVHGQIQH PVQPFLRQGF GYALGLLRRF DVGGRVGAHQ
151  PAFDQPGAIL PPRRQLARQR PTVQTALRQP PQRRRKIAPR QVLRHAACIF
201  RRHLCQQCKQ FFQIAPVCRN RVLRLALAHD VFQISVKIRR KPVQNHNRPT
251  QISKNQ*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 111>:

m032.seq (partial)

```
  1  ATGCGGCGAA ACGTGCmTGC mGTCGCCGTT kTGCGCCGCC CATTGCGCCA
 51  AACGTTTTTG GATTTGGCGT TGGCTCAGGC GCGTGCCGTT CCTGCCGGTA
101  AACAGGGCTT TGCCGTCCGA TGCCGTCTGA CGCAGCGGCA GATAGTTTTT
151  CAGGGCTTCC ACGCTTTTGC CGACCAGCGG CACCTGCCGC TgTT.GCGCC
201  CTTTGCCGAT AAcGTGTACC CACGCyTCGT CCAAATAGAC ATCATCTGCA
251  TTCAAGCCGT GTATCTCGCT CACGCGCAAA CCGCTGCCGT ACATCAGTTC
301  GAACAGGGCG TGGTCGCGCA CCGCCAGCGG GTCGCCGCCG TCCACGGGCA
351  AATCCAGCAT CCGGTTCACC CATTCCTGCG GCAGGGCTTT GGGTACGCGC
401  TCGGGCTGCT TCGGCGGTTT GATGTCGGCG GTCGGGTCGG CGTGCATCAG
451  GCCGCGCTTT ACCAGCCAAA CGCAATACTG CCGCCAAGAC GAAAGCTTGC
501  GAGCCAGCGT CCGTTCCCCC AAACCGCG...
```

This corresponds to the amino acid sequence <SEQ ID 112; ORF 032>:

m032.pep (partial)

```
  1  MRRNVXAVAV XRRPLRQTFL DLALAQARAV PAGKQGFAVR CRLTQRQIVF
 51  QGFHAFADQR HLPLXAPFAD NVYPRXVQID IICIQAVYLA HAQTAAVHQF
101  EQGVVAHRQR VAAVHGQIQH PVQPFLRQGF GYALGLLRRF DVGGRVGVHQ
151  AALYQPNAIL PPRRKLASQR PFPQTA...
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 113>:

a032.seq

```
  1  ATGCGGCGAA ACGTGCCTGC CGTCGCCGTT TTGCGCCGCC CATTGCGCCA
 51  AACGTTTTTG GATTTGGCGT TGGCTCAGGC GCGTGCCGTT CCTGCCGGTA
101  AACAGGGCTT TGCCGTCCGA TGCCGTCTGA CGCAGCGGCA GATAGTTTTT
151  CAGGGCTTCC ACGCTTTTGC CGGTCAGCGG AACCTGCCGC TGCTTGCGTC
201  CTTTGCCGGT AACGTGTACC CACGCCTCGT CCAAATATAC ATCATCTGCA
251  TTCAAGCCGT GTATCTCGCT CACGCGCAAA CCGCTGCCGT ACATCAGTTC
301  GAACAGCGCG TGATCGCGCA CCGCCAGCGG GTCGCCGCCG TCCACGGGCA
351  AATCCAGCAT CCGGTTCAGC CATTCCTGCG GCAGGGCTTT GGGTACGCGC
401  TCGGGCTGCT TCGGCGGTTT GATGTCGGCG GTCGGGTCGG TATGCAGCAG
451  ACCGCGTTTG ACCAGCCAGG CGCAATACTG CCGCCAAGAC GACAGCTTGC
501  GCGCCAGCGT CCGCGCATTC AAACCGCGCT GCGACAGCCG CCGCAACGCC
551  GCCGTAAAAT CGCGCTGCGA CAAGCCCTGC GGCACGCCGC CTGCATCTTC
601  AGACGGCATT TGTGCCAACA GCGCAAACAG TTCTTCCAAA TCGCGCCGGT
651  ATGCCGCCAC CGTGTGCTCC GACTTGCCCT CGCGCACGAT GTTTTCCAAA
701  TAAGCGTCAA AATGCGCCGC AAACCCGTCC AAAACCATAA CCGCCCCACA
751  CAAATATCAA AAAAACAGTG A
```

This corresponds to the amino acid sequence <SEQ ID 114; ORF 032.a>:

a032.pep

```
  1  MRRNVPAVAV LRRPLRQTFL DLALAQARAV PAGKQGFAVR CRLTQRQIVF
 51  QGFHAFAGQR NLPLLASFAG NVYPRLVQIY IICIQAVYLA HAQTAAVHQF
101  EQRVIAHRQR VAAVHGQIQH PVQPFLRQGF GYALGLLRRF DVGGRVGMQQ
151  TAFDQPGAIL PPRRQLARQR PRIQTALRQP PQRRRKIALR QALRHAACIF
201  RRHLCQQRKQ FFQIAPVCRH RVLRLALAHD VFQISVKMRR KPVQNHNRPT
251  QISKKQ*
```

m032/a032 88.1% identity over a 176 aa overlap

```
                    10        20        30        40        50        60
m032.pep    MRRNVXAVAVXRRPLRQTFLDLALAQARAVPAGKQGFAVRCRLTQRQIVFQGFHAFADQR
            ||||| |||| |||||||||||||||||||||||||||||||||||||||||||||| ||
a032        MRRNVPAVAVLRRPLRQTFLDLALAQARAVPAGKQGFAVRCRLTQRQIVFQGFHAFAGQR
                    10        20        30        40        50        60
```

```
                    70        80        90       100       110       120
m032.pep    HLPLXAPFADNVYPRXVQIDIICIQAVYLAHAQTAAVHQFEQGVVAHRQRVAAVHGQIQH
            :||| | || ||||| ||| |||||||||||||||||||||| |:|||||||||||||||
a032        NLPLLASFAGNVYPRLVQIYIICIQAVYLAHAQTAAVHQFEQRVIAHRQRVAAVHGQIQH
                    70        80        90       100       110       120

                   130       140       150       160       170
m032.pep    PVQPFLRQGFGYALGLLRRFDVGGRVGVHQAALYQPNAILPPRRKLASQRPFPQTA
            |||||||||||||||||||||||||||::|:|: ||:|||||||:|| |||  |||
a032        PVQPFLRQGFGYALGLLRRFDVGGRVGMQQTAFDQPGAILPPRRQLARQRPRIQTALRQP
                   130       140       150       160       170       180

a032        PQRRRKIALRQALRHAACIFRRHLCQQRKQFFQIAPVCRHRVLRLALAHDVFQISVKMRR
                   190       200       210       220       230       240
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 032 shows 86.4% identity over a 176 aa overlap with a predicted ORF (ORF 032.ng) from *N. gonorrhoeae:*

m032/g032

```
                    10        20        30        40        50        60
m032.pep    MRRNVXAVAVXRRPLRQTFLDLALAQARAVPAGKQGFAVRCRLTQRQIVFQGFHAFADQR
            ||||| |||| |||  ::||||||||||||||||||||||||||||||||||||||| ||
g032        MRRNVPAVAVLRRPRFEAFLDLALAQARAVPAGKQGFAVRCRLTQRQIVFQGFHAFAGQR
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m032.pep    HLPLXAPFADNVYPRXVQIDIICIQAVYLAHAQTAAVHQFEQGVVAHRQRVAAVHGQIQH
            :| | |||| ||||| ||| |||||||||||||||||||:|| |||||||||||||||||
g032        NLTLLAPFAGNVYPRFVQIYIICIQAVYLAHAQTAAVHQLEQRVVAHRQRVAAVHGQIQH
                    70        80        90       100       110       120

                   130       140       150       160       170
m032.pep    PVQPFLRQGFGYALGLLRRFDVGGRVGVHQAALYQPNAILPPRRKLASQRPFPQTA
            |||||||||||||||||||||||||||:|| |: ||:|||||||:|| |||  |||
g032        PVQPFLRQGFGYALGLLRRFDVGGRVGAHQPAFDQPGAILPPRRQLARQRPTVQTALRQP
                   130       140       150       160       170       180

g032        PQRRRKIAPRQVLRHAACIFRRHLCQQCKQFFQIAPVCRNRVLRLALAHDVFQISVKIRR
                   190       200       210       220       230       240
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 115>:

g033.seq

```
  1  ATGGCGGCGG CGGACAAACT CTTGGGCGGC GACCGCCGCA GCGTCGCCAT
 51  CATCGGAGAC GGCGCGATGA CGGCGGGGCA GGCGTTTGAA GCCTTGAATT
101  GCGCGGGCGA TATGGATGTG GATTTGCTGG TCGTCCTCAA CGACAACGAA
151  ATGTCGATTT CCCCCAACGT CGGCGCGTTG CCCAAATATC TTGCCAGCAA
201  CGTCGTGCGC GATATGCACG GACTGTTGAG TACCGTCAAA GCGCAAAcgg
251  GCAAGGTATT AGACAAAATA CCCGGCGCGA TGGagtTTGC CCAAAAAGTC
301  GAACAcaaaA TCAAAACCCT TGCCGAAGAA GCCGAACACG CCAAACAGTC
351  GCTGTCGCTG TTTGAAAATT TCGGCTTCCG CTACACCGGC CCCGTGGACG
401  GACACAACGT CGAGAATCTG GTGGACGTAT TGAAAGACTT GCGCAGCCGC
451  AAAGGCCCTC AGTTGCTGCA CGTCATCACC AAAAAGGGCA ACGGCTACAA
501  ACTCGCCGAA AACGACCCCg tcaAATACCA CGCCGTCGCc aACCTGCcta
551  AAGAAGGCGG GGCGCAAATg ccGTCTGAAA AAGAACCCAA GCCCGCCgCc
601  aaaccgACCT ATACCCAAGT ATTCGGCAAA TGGCTGTGCG ACCGGGCGGC
651  GGCAGATTCC CGACTGGTTG CGATTACCCC CGCCATGCGC GAGGGCAGCG
```

```
 701  GACTGGTGGA GTTTGAACAA CGATTCCCCG ACCGCTATTT CGATGTCGGC
 751  ATCGCCGAGC AGCACGCCGT tacCTTTGCC GGCGGTTTGG CGTGCGAAGG
 801  CATGAAGCCC GTCGTGGCGA TTTATTCCAC CTTTTTACAA CGCGCCTACG
 851  ACCAACTGGT GCACGACATC GCCCTGCAAA ACCTGCCCGT TTTGTTTGCC
 901  GTCGACCGTG CGGGCATCGT CGGCGCGGAC GGTCCGACCC ATGCCGGCTT
 951  GTACGATTTG AGCTTCTTGC GCTGTGTGCC GAACATGATT GTTGCCGCGC
1001  CGAGCGATGA AAACGAATGC CGCCTGCTGC TTTCGACCTG CTATCAGGCG
1051  GATGCGCCCG CCGCCGTCCG CTATCCGCGC GGCACGGGTA CGGGCGCGCC
1101  GGTTTCAGAC GGCATGGAAA CCGTGGAAAT CGGCAAGGGC ATTATCCGCC
1151  GCGAAGGTGA GAAAACCGCC TTcatTGCCT TCGGCAGTAT GGTCGCCACC
1201  GCATTGGCGG TTGCCGAAAA ACTGAACGCC ACCGTCGCCG ATATGCGCTt
1251  cgtcaaacCG ATAGACGAAG AGTTGATTGT CCGCCTTGCC CGAAGCCAcg
1301  accGCATCGT TACCCTTGAA GAAAACGCCG AACAGGGCGG CGCAGGCGGC
1351  GCGGTCTTGG AAGTGTTGGC GAAACACGGC ATCTGCAAAC CCGTTTTGCT
1401  TTTGGGCGTT GCCGATACCG TAACCGAACA CGGCGATCCG AAAAAACTTT
1451  TGGACGATTT GGGTTTGAGT GCCGAAGCGG TGGAACGCCG GGTGCGCGAG
1501  TGGCTGCCGG ACCGTGATGC GGCAAATTAA
```

This corresponds to the amino acid sequence <SEQ ID 116; ORF 033.ng>:

g033.pep

```
  1  MAAADKLLGG DRRSVAIIGD GAMTAGQAFE ALNCAGDMDV DLLVVLNDNE
 51  MSISPNVGAL PKYLASNVVR DMHGLLSTVK AQTGKVLDKI PGAMEFAQKV
101  EHKIKTLAEE AEHAKQSLSL FENFGFRYTG PVDGHNVENL VDVLKDLRSR
151  KGPQLLHVIT KKGNGYKLAE NDPVKYHAVA NLPKEGGAQM PSEKEPKPAA
201  KPTYTQVFGK WLCDRAAADS RLVAITPAMR EGSGLVEFEQ RFPDRYFDVG
251  IAEQHAVTFA GGLACEGMKP VVAIYSTFLQ RAYDQLVHDI ALQNLPVLFA
301  VDRAGIVGAD GPTHAGLYDL SFLRCVPNMI VAAPSDENEC RLLLSTCYQA
351  DAPAAVRYPR GTGTGAPVSD GMETVEIGKG IIRREGEKTA FIAFGSMVAT
401  ALAVAEKLNA TVADMRFVKP IDEELIVRLA RSHDRIVTLE ENAEQGGAGG
451  AVLEVLAKHG ICKPVLLLGV ADTVTEHGDP KKLLDDLGLS AEAVERRVRE
501  WLPDRDAAN*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 117>:

m033.seq

```
   1  ATGGCGGCGG CAGACAAACT CTTGGGCAGC GACCGCCGCA GCGTCGCCAT
  51  CATCGGCGAC GGCGCGATGA CGGCGGGGCA GGCGTTTGAA GCCTTGAATT
 101  GCGCaG.CGA TATGGATGTr GATTTGCTrG TCGTCCTCAA CGACAACGAA
 151  ATGTCGATTT CCCCCAACGT CGGCGCGCTG CCGAAATACC TTGCCAGCAA
 201  CGTCGTGCGC GATATGCACG GCCTGTTGAG TACCGTCAAA GCGCAAACGG
 251  GCAAGGTATT AGACAAAATA CCCGGCGCGA TGGAGTTTGC CCAAAAAGTC
 301  GAACACAAAA TCAAAACCCT TGCCGAAGAA GCCGAACACG CCAAACAGTC
 351  GCTGTCTTTG TTTGAAAACT TCGGCTTCCG CTACACCGGC CCCGTGGACG
 401  GACACAACGT CGAAAATCTG GTGGACGTAT TGAAAGACTT GCGCAGCCGC
 451  AAAGGCCCTC AGTTGCTGCA CGTCATCACC AAAAAGGGCA ACGGCTACAA
 501  ACTCGCCGAA AACGACCCCG TCAAATACCA CGCCGTCGCC AACCTGCCTA
 551  AAGAAAGCGC GGCGCAAATG CCGTCTGAAA AAGAACCCAA GCCCGCCGCC
 601  AAACCGACCT ATACCCAAGT GTTCGGCAAA TGGCTGTGCG ACCGGGCGGC
 651  GGCAGATTCC CGACTGGTTG CGATTACCCC CGCCATGCGC GAGGGCAGCG
 701  GCTTGGTTGA GTTTGAACAA CGATTCCCCG ACCGCTATTT CGATGTCGGC
 751  ATCGCCGAGC AGCACGCCGT TACCTTTGCC GGCGGTTTGG CTTGCGAAGG
 801  GATGAAGCCC GTCGTGGCGA TTTATTCCAC CTTTTTACAA CGCGCCTACG
 851  ACCAACTGGT GCACGACATC GCCCTGCAAA ACCTACCCGT TTTGTTTGCC
 901  GTCGACCGCG CGGGCATCGT CGGCGCGGAC GGCCCGACCC ATGCCGGTCT
 951  GTACGATTTG AGCTTTTTGC GCTGCGTGCC GAACATGATT GTCGCCGCGC
1001  CGAGCGATGA AAACGAATGC CGCCTGTTGC TTTCGACCTG CTATCAGGCA
1051  GACGCGCCCG CCGCCGTCCG CTATCCGCGC GGCACGGGTA CGGGCGCGCC
1101  GGTTTCAGAC GGCATGGAAA CCGTGGAAAT CGGCAAGGGC ATTATCCGCC
1151  GCGAAGGTGA GAAAACCGCA TTCATTGCCT TCGGCAGTAT GGTCGCCCCC
1201  GCATTGGCGG TTGCCGAAAA ACTGAACGCC ACCGTCGCCG ATATGCGCTT
1251  CGTCAAACCG ATAGACGAAG AGTTGATTGT CCGCCTTGCC CGAAGCCACG
```

```
1301  ACCGCATCGT TACCCTTGAA GAAAACGCCG AACAGGGCGG CGCAGGCGGC
1351  GCGGTGCTGG AAGTATTGGC GAAACACGGC ATCTGCAAAC CCGTTTTGCT
1401  TTTGGGCGTT GCCGATACCG TAACCGGACA CGGCGATCCG AAAAAACTTT
1451  TAGACGATTT GGGCTTGAGT GCCGAAGCGG TGGAACGGCG TGTGCGCGCG
1501  TGGCTGTCGG ATCGGGATGC GGCAAATTAA
```

This corresponds to the amino acid sequence <SEQ ID 118; ORF 033>:

```
m033.pep
  1  MAAADKLLGS DRRSVAIIGD GAMTAGQAFE ALNCAXDMDV DLLVVLNDNE
 51  MSISPNVGAL PKYLASNVVR DMHGLLSTVK AQTGKVLDKI PGAMEFAQKV
101  EHKIKTLAEE AEHAKQSLSL FENFGFRYTG PVDGHNVENL VDVLKDLRSR
151  KGPQLLHVIT KKGNGYKLAE NDPVKYHAVA NLPKESAAQM PSEKEPKPAA
201  KPTYTQVFGK WLCDRAAADS RLVAITPAMR EGSGLVEFEQ RFPDRYFDVG
251  IAEQHAVTFA GGLACEGMKP VVAIYSTFLQ RAYDQLVHDI ALQNLPVLFA
301  VDRAGIVGAD GPTHAGLYDL SFLRCVPNMI VAAPSDENEC RLLLSTCYQA
351  DAPAAVRYPR GTGTGAPVSD GMETVEIGKG IIRREGEKTA FIAFGSMVAP
401  ALAVAEKLNA TVADMRFVKP IDEELIVRLA RSHDRIVTLE ENAEQGGAGG
451  AVLEVLAKHG ICKPVLLLGV ADTVTGHGDP KKLLDDLGLS AEAVERRVRA
501  WLSDRDAAN*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 119>:

```
a033.seq
   1  ATGGCGGCGG CGGACAAACA GTTGGGCAGC GACCGCCGCA GCGTCGCCAT
  51  CATCGGCGAC GGCGCGATGA CGGCGGGTCA GGCGTTTGAA GCCTTGAACT
 101  GCGCGGGCGA TATGGATGTG GATTTGCTGG TCGTCCTCAA CGACAACGAA
 151  ATGTCGATTT CCCCCAACGT CGGTGCGTTG CCCAAATACC TTGCCAGCAA
 201  CGTCGTGCGC GATATGCACG GACTGTTGAG TACCGTCAAA GCGCAAACGG
 251  GCAAGGTATT AGACAAAATA CCCGGCGCGA TGGAGTTTGC CCAAAAAGTC
 301  GAACATAAAA TCAAAACCCT TGCCGAAGAA GCCGAACACG CCAAACAGTC
 351  ACTGTCTTTG TTTGAAAACT TCGGCTTCCG CTATACCGGC CCCGTGGACG
 401  GACACAACGT CGAAAATCTG GTCGATGTAT TGGAAGACCT GCGCGGACGC
 451  AAAGGCCCGC AGCTTCTGCA CGTCATCACC AAAAAGGGCA ACGGCTACAA
 501  ACTCGCCGAA AACGATCCCG TCAAATACCA CGCCGTCGCC AACCTGCCTA
 551  AAGAAAGCGC GGCGCAAATG CCGTCTGAAA AAGAACCCAA GCCCGCCGCC
 601  AAACCGACCT ATACCCAAGT GTTCGGCAAA TGGCTGTGCG ACCGGGCGGC
 651  GGCAGATTCC CGACTGGTTG CGATTACCCC CGCCATGCGC GAGGGCAGCG
 701  GCTTGGTTGA GTTTGAACAA CGATTCCCCG ACCGCTATTT CGATGTCGGC
 751  ATCGCCGAGC AGCACGCCGT TACCTTTGCC GGCGGTTTGG CTTGCGAAGG
 801  GATGAAGCCC GTCGTGGCGA TTTATTCCAC CTTTTTACAA CGCGCCTACG
 851  ACCAACTGGT GCACGACATC GCCCTGCAAA ACCTGCCCGT TTTGTTTGCC
 901  GTCGACCGCG CGGGCATCGT CGGCGCGGAC GGCCCGACCC ATGCCGGTTT
 951  GTACGATTTA AGCTTTTTGC GCTGCATTCC GAATATGATT GTCGCCGCGC
1001  CGAGCGATGA AAATGAATGC CGCCTGCTGC TTTCGACCTG CTATCAGGCA
1051  GACGCGCCCG CCGCCGTCCG CTATCCGCGC GGCACGGGTA CGGGCGTGCC
1101  GGTTTCAGAC GGCATGGAAA CCGTGGAAAT CGGCAAGGGC ATTATCCGCC
1151  GCGAAGGTGA GAAAACCGCA TTCATTGCCT TCGGCAGTAT GGTCGCCCCT
1201  GCATTGGCGG TCGCCGGAAA ACTGAACGCC ACCGTCGCCG ATATGCGCTT
1251  CGTCAAACCG ATAGACGAAG AGTTGATTGT CCGCCTTGCC CGAAGCCACG
1301  ACCGCATCGT TACCCTTGAA GAAAACGCCG AACAGGGCGG CGCAGGCAGC
1351  GCGGTGCTGG AAGTGTTGGC GAAACACGGC ATCTGCAAAC CCGTCTTGCT
1401  TTTGGGCGTT GCCGATACCG TAACCGGACA CGGCGATCCG AAAAAACTTT
1451  TAGACGATTT GGGCTTGAGT GCCGAAGCGG TGGAACGGCG TGTGCGCGCG
1501  TGGCTGTCGG ATCGGGATGC GGCAAATTAA
```

This corresponds to the amino acid sequence <SEQ ID 120; ORF 033.a>:

```
a033.pep
  1  MAAADKQLGS DRRSVAIIGD GAMTAGQAFE ALNCAGDMDV DLLVVLNDNE
 51  MSISPNVGAL PKYLASNVVR DMHGLLSTVK AQTGKVLDKI PGAMEFAQKV
101  EHKIKTLAEE AEHAKQSLSL FENFGFRYTG PVDGHNVENL VDVLEDLRGR
151  KGPQLLHVIT KKGNGYKLAE NDPVKYHAVA NLPKESAAQM PSEKEPKPAA
201  KPTYTQVFGK WLCDRAAADS RLVAITPAMR EGSGLVEFEQ RFPDRYFDVG
251  IAEQHAVTFA GGLACEGMKP VVAIYSTFLQ RAYDQLVHDI ALQNLPVLFA
```

```
301  VDRAGIVGAD GPTHAGLYDL SFLRCIPNMI VAAPSDENEC RLLLSTCYQA
351  DAPAAVRYPR GTGTGVPVSD GMETVEIGKG IIRREGEKTA FIAFGSMVAP
401  ALAVAGKLNA TVADMRFVKP IDEELIVRLA RSHDRIVTLE ENAEQGGAGS
451  AVLEVLAKHG ICKPVLLLGV ADTVTGHGDP KKLLDDLGLS AEAVERRVRA
501  WLSDRDAAN*
```

m033/a033 98.4% identity over a 509 aa overlap

```
                    10        20        30        40        50        60
m033.pep    MAAADKLLGSDRRSVAIIGDGAMTAGQAFEALNCAXDMDVDLLVVLNDNEMSISPNVGAL
            |||||| |||||||||||||||||||||||    .|| |||||||||||||||||||||||
a033        MAAADKQLGSDRRSVAIIGDGAMTAGQAFEALNCAGDMDVDLLVVLNDNEMSISPNVGAL
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m033.pep    PKYLASNVVRDMHGLLSTVKAQTGKVLDKIPGAMEFAQKVEHKIKTLAEEAEHAKQSLSL
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a033        PKYLASNVVRDMHGLLSTVKAQTGKVLDKIPGAMEFAQKVEHKIKTLAEEAEHAKQSLSL
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m033.pep    FENFGFRYTGPVDGHNVENLVDVLKDLRSRKGPQLLHVITKKGNGYKLAENDPVKYHAVA
            ||||||||||||||||||||||||:|||:|||||||||||||||||||||||||||||||
a033        FENFGFRYTGPVDGHNVENLVDVLEDLRGRKGPQLLHVITKKGNGYKLAENDPVKYHAVA
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m033.pep    NLPKESAAQMPSEKEPKPAAKPTYTQVFGKWLCDRAAADSRLVAITPAMREGSGLVEFEQ
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a033        NLPKESAAQMPSEKEPKPAAKPTYTQVFGKWLCDRAAADSRLVAITPAMREGSGLVEFEQ
                   190       200       210       220       230       240

                   250       260       270       280       290       300
m033.pep    RFPDRYFDVGIAEQHAVTFAGGLACEGMKPVVAIYSTFLQRAYDQLVHDIALQNLPVLFA
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a033        RFPDRYFDVGIAEQHAVTFAGGLACEGMKPVVAIYSTFLQRAYDQLVHDIALQNLPVLFA
                   250       260       270       280       290       300

                   310       320       330       340       350       360
m033.pep    VDRAGIVGADGPTHAGLYDLSFLRCVPNMIVAAPSDENECRLLLSTCYQADAPAAVRYPR
            |||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
a033        VDRAGIVGADGPTHAGLYDLSFLRCIPNMIVAAPSDENECRLLLSTCYQADAPAAVRYPR
                   310       320       330       340       350       360

                   370       380       390       400       410       420
m033.pep    GTGTGAPVSDGMETVEIGKGIIRREGEKTAFIAFGSMVAPALAVAEKLNATVADMRFVKP
            |||||:||||||||||||||||||||||||||||||||||||||||| ||||||||||||
a033        GTGTGVPVSDGMETVEIGKGIIRREGEKTAFIAFGSMVAPALAVAGKLNATVADMRFVKP
                   370       380       390       400       410       420

                   430       440       450       460       470       480
m033.pep    IDEELIVRLARSHDRIVTLEENAEQGGAGGAVLEVLAKHGICKPVLLLGVADTVTGHGDP
            |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
a033        IDEELIVRLARSHDRIVTLEENAEQGGAGSAVLEVLAKHGICKPVLLLGVADTVTGHGDP
                   430       440       450       460       470       480

                   490       500       510
m033.pep    KKLLDDLGLSAEAVERRVRAWLSDRDAANX
            ||||||||||||||||||||||||||||||
a033        KKLLDDLGLSAEAVERRVRAWLSDRDAANX
                   490       500       510
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 033 shows 98.4% identity over a 509 aa overlap with a predicted ORF (ORF 033.ng) from *N. gonorrhoeae:*

m033/g033

```
m033.pep    MAAADKLLGSDRRSVAIIGDGAMTAGQAFEALNCAXDMDVDLLVVLNDNEMSISPNVGAL   60
            |||||||||:||||||||||||||||||||||||| ||||||||||||||||||||||||
g033        MAAADKLLGGDRRSVAIIGDGAMTAGQAFEALNCAGDMDVDLLVVLNDNEMSISPNVGAL   60

m033.pep    PKYLASNVVRDMHGLLSTVKAQTGKVLDKIPGAMEFAQKVEHKIKTLAEEAEHAKQSLSL  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g033        PKYLASNVVRDMHGLLSTVKAQTGKVLDKIPGAMEFAQKVEHKIKTLAEEAEHAKQSLSL  120

m033.pep    FENFGFRYTGPVDGHNVENLVDVLKDLRSRKGPQLLHVITKKGNGYKLAENDPVKYHAVA  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g033        FENFGFRYTGPVDGHNVENLVDVLKDLRSRKGPQLLHVITKKGNGYKLAENDPVKYHAVA  180

m033.pep    NLPKESAAQMPSEKEPKPAAKPTYTQVFGKWLCDRAAADSRLVAITPAMREGSGLVEFEQ  240
            |||||::|||||||||||||||||||||||||||||||||||||||||||||||||||||
g033        NLPKEGGAQMPSEKEPKPAAKPTYTQVFGKWLCDRAAADSRLVAITPAMREGSGLVEFEQ  240

m033.pep    RFPDRYFDVGIAEQHAVTFAGGLACEGMKPVVAIYSTFLQRAYDQLVHDIALQNLPVLFA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g033        RFPDRYFDVGIAEQHAVTFAGGLACEGMKPVVAIYSTFLQRAYDQLVHDIALQNLPVLFA  300

m033.pep    VDRAGIVGADGPTHAGLYDLSFLRCVPNMIVAAPSDENECRLLLSTCYQADAPAAVRYPR  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g033        VDRAGIVGADGPTHAGLYDLSFLRCVPNMIVAAPSDENECRLLLSTCYQADAPAAVRYPR  360

m033.pep    GTGTGAPVSDGMETVEIGKGIIRREGEKTAFIAFGSMVAPALAVAEKLNATVADMRFVKP  420
            ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
g033        GTGTGAPVSDGMETVEIGKGIIRREGEKTAFIAFGSMVATALAVAEKLNATVADMRFVKP  420

m033.pep    IDEELIVRLARSHDRIVTLEENAEQGGAGGAVLEVLAKHGICKPVLLLGVADTVTGHGDP  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
g033        IDEELIVRLARSHDRIVTLEENAEQGGAGGAVLEVLAKHGICKPVLLLGVADTVTEHGDP  480

m033.pep    KKLLDDLGLSAEAVERRVRAWLSDRDAANX  510
            ||||||||||||||||||| || |||||||
g033        KKLLDDLGLSAEAVERRVREWLPDRDAANX  510
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 121>:

g034.seq

```
  1  ATGAGCCGTT TATGGTTTTT TGCCGTAAAA AACATTATAA TCCGCCTTAT
 51  TTACCTATTG CCCAAGGAGA CACAAATGGC ACTCGTATCC ATGCGCCAAC
101  TGCTTGACCA CGCCGCCGAA AACAGCTACG GCCTGCCCGC GTTCAACGTC
151  AACAACCTCG AACAAATGCG CGCCATTATG GAAGCCGCCG ACCAAGTCAA
201  CGCGCCCGTC ATCGTACAGG CGAGCGCAGG TGCGCGCAAA TACGcggGCG
251  CGCCGTTTTT GCGCCACCTG ATTCTGGCGG CAGTCGAAGA ATTTCCGCAC
301  ATCCCCGTCG TGATGCACCA AGACCACGGC GCATCGCCCG ACGTgtgCCA
351  ACGCTCCATC CAACTGGGCT TCTCCTCCGT GATGATGGAC GGCTCTTTGC
401  TCGAAGACGG CAAAACCCCT TCTTCTTACG AATACAACGT CAACGCCACC
451  CGTACCGTCG TCAACTTCTC CCACGCCTGC GGCGTGTCCG TCGAAGGCGA
501  AATCGGCGTA TTGGGCAACC TCGAAACCGG CGAAGCAGGC GAAGAAGACG
551  GAGTGGGCGC GGCAGGCAAA CTCTCACACG ACCAAATGCT CACCAGCGTT
601  GAAGATGCCG TGCGTTTCGT TAAAGATACC GGCGTTGACG CATTGGCGAT
651  TGCCGTCGGC ACCAGCCACG GCGCATACAA ATTCACCCGT CCGCCCACAG
```

```
 701  GCGACGTATT GCGTATCGAC CGCATCAAGG AAATCCACCA AGCCCTGCCC
 751  AATACACACA TCGTGATGCA CGgctCCAGC TCCGTTCCGC AAGAatgGCT
 801  GAAAGTCATC AACGAATACG GCGGCAATAT CGGCGAAACC TACGGCGTGC
 851  CGGTTGAAGA AATCGTCGAA GGCATCAAAC ACGGCGTGCG CAAAGTCAAC
 901  ATCGATACCG ACCTGCGCCT CGCTTCCACC GGCGCGGTAC GCCGCTACCT
 951  TGCCGAAAAC CCGTCCGACT TTGATCCGCG CAAATACTTG GGCAAAACCA
1001  TTGAAGCGAT GAAGCAAATC TGCCTCGACC GTTATCTTGC GTTCGGTTGC
1051  GAAGGTCAGG CAGGCAAAAT CAAACCTGTT TCGTTGGAAA AAATGGCAAG
1101  CCGTTATGCC AAGGGCGAAT TGAACCAAAT CGTCAAATAA
```

This corresponds to the amino acid sequence <SEQ ID 122; ORF 034.ng>:

g034.pep

```
   1  MSRLWFFAVK NIIIRLIYLL PKETQMALVS MRQLLDHAAE NSYGLPAFNV
  51  NNLEQMRAIM EAADQVNAPV IVQASAGARK YAGAPFLRHL ILAAVEEFPH
 101  IPVVMHQDHG ASPDVCQRSI QLGFSSVMMD GSLLEDGKTP SSYEYNVNAT
 151  RTVVNFSHAC GVSVEGEIGV LGNLETGEAG EEDGVGAAGK LSHDQMLTSV
 201  EDAVRFVKDT GVDALAIAVG TSHGAYKFTR PPTGDVLRID RIKEIHQALP
 251  NTHIVMHGSS SVPQEWLKVI NEYGGNIGET YGVPVEEIVE GIKHGVRKVN
 301  IDTDLRLAST GAVRRYLAEN PSDFDPRKYL GKTIEAMKQI CLDRYLAFGC
 351  EGQAGKIKPV SLEKMASRYA KGELNQIVK*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 123>:

m034.seq (partial)

```
   1  ATGAGCTGTT TATGGTTTTT TGCTGTAAAA AACATTATAA TCCGCCTTAT
  51  TTACCTATTG CCCAAGGAGA CACAAATGGC ACTCGTATCC ATGCGCCAAC
 101  TGCTTGATCA TGCTGCCGAA wACAGCTACG GCyTGCCGGC GTTCAACGTC
 151  AACAACCTCG wACAGATGCG CGCCATCATG GAGGCTGCAG ACCAAGTCGA
 201  CGCCCCCGTC ATCGTACAGG CGAGTGCCGG TGCGCGCAAA TATGCGGGTG
 251  CGCCGTTTTT ACGCCACCTG ATTTTGGCGG CTGTCGAAGT ATTTCCACAC
 301  ATCCCCGTCG TCATGCACCA AGACCACGGC GCATCACCCG ACGTGTGCCA
 351  ACGCTCCATC CAACTGGGCT TCTCCTCTGT AATGATGGAC GGCTCGCTGA
 401  TGGAAGACGG CAAAACCCCT TCTTCTTACG AATACAACGT CAACGCCACA
 451  CGTACCGTGG TTAACTTCTC CCACGCTTGC GGCGTATCCG TTGAAGGCGA
 501  AATCGGCGTA TTGGGCAACC TCGAAACCGG CGATGCAGGC GAAGAAGACG
 551  GTGTAGGCGC AGTGGGCAAA CTTTCCCACG ACCAAATGCT GACCAGCGTC
 601  GAAGATGCCG TATGTTTCGT TAAAGATACC GGCGTTGACG CATTGGCTAT
 651  TGCCGTCGGC ACCAGCCACG GCGCATACAA ATTCACCCGT CCGCCCACAG
 701  GCGATGTATT ACGTATCGAC CGCATCAAAG AAATCCACCA AGCCCTGCCC
 751  AATACACACA TCGTGATGCA C...
```

This corresponds to the amino acid sequence <SEQ ID 124; ORF 034>:

m034.pep (partial)

```
   1  MSCLWFFAVK NIIIRLIYLL PKETQMALVS MRQLLDHAAE XSYGLPAFNV
  51  NNLXQMRAIM EAADQVDAPV IVQASAGARK YAGAPFLRHL ILAAVEVFPH
 101  IPVVMHQDHG ASPDVCQRSI QLGFSSVMMD GSLMEDGKTP SSYEYNVNAT
 151  RTVVNFSHAC GVSVEGEIGV LGNLETGDAG EEDGVGAVGK LSHDQMLTSV
 201  EDAVCFVKDT GVDALAIAVG TSHGAYKFTR PPTGDVLRID RIKEIHQALP
 251  NTHIVMH...
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 125>:

a034.seq

```
   1  ATGAGCCGTT TATGGTTTTT TGCCGCAAAA AACATTATAA TCCGCCTTAT
  51  TTACCTATTG CCCAAGGAGA CACAAATGGC ACTCGTATCC ATGCGCCAAC
 101  TGCTTGATCA TGCTGCCGAA AACAGCTACG GCCTGCCCGC GTTCAACGTC
 151  AACAACCTCG AACAAATGCG CGCCATTATG GAAGCCGCCG ACCAAGTCAA
 201  CGCGCCCGTC ATCGTACAGG CGAGCGCAGG TGCGCGCAAA TACGCGGGCG
 251  CGCCGTTTTT GCGCCACCTG ATTTTGGCGG CTGTCGAAGA ATTTCCGCAC
 301  ATCCCCGTCG TGATGCACCA AGACCACGGC GCATCGCCCG ACGTGTGCCA
 351  ACGCTCCATC CAACTGGGCT TTTCCTCCGT GATGATGGAC GGCTCGCTGA
 401  TGGAAGACGG CAAAACCCCT TCTTCTTATG AATACAACGT CAACGCCACC
```

```
 451  CGTACCGTGG TTAATTTCTC CCACGCCTGC GGCGTATCCG TTGAAGGCGA
 501  AATCGGCGTA TTGGGCAACC TCGAAACTGG CGAAGCCGGC GAAGAAGACG
 551  GTGTAGGCGC AGTGGGCAAA CTTTCCCACG ACCAAATGCT CACCAGCGTC
 601  GAAGATGCCG TGCGTTTCGT TAAAGATACC GGCGTTGACG CATTGGCGAT
 651  TGCCGTCGGC ACCAGCCACG GCGCGTACAA ATTCACCCGT CCGCCCACAG
 701  GCGACGTGTT GCGTATCGAC CGCATCAAAG AAATCCACCA AGCCCTGCCC
 751  AATACACACA TCGTGATGCA CGGCTCCAGC TCCGTTCCGC AAGAATGGCT
 801  GAAAGTCATC AACGAATACG GCGGCAATAT CGGCGAAACC TACGGCGTGC
 851  CGGTTGAAGA AATCGTCGAA GGCATCAAAC ACGGCGTGCG TAAAGTCAAC
 901  ATCGATACCG ACTTGCGCCT TGCTTCCACC GGCGCGGTAC GCCGCTACCT
 951  TGCCGAAAAC CCGTCCGACT TCGATCCGCG CAAATATTTG AGCAAAACCA
1001  TTGAAGCGAT GAAGCAAATC TGCCTCGACC GCTACCTCGC GTTCGGTTGC
1051  GAAGGTCAGG CAGGCAAAAT CAAACCGGTT TCCTTGGAAA AAATGGCAAA
1101  CCGTTATGCC AAGGGCGAAT TGAACCAAAT CGTCAAATAA
```

This corresponds to the amino acid sequence <SEQ ID 126; ORF 034.a>:

```
a034.pep
  1  MSRLWFFAAK NIIIRLIYLL PKETQMALVS MRQLLDHAAE NSYGLPAFNV
 51  NNLEQMRAIM EAADQVNAPV IVQASAGARK YAGAPFLRHL ILAAVEEFPH
101  IPVVMHQDHG ASPDVCQRSI QLGFSSVMMD GSLMEDGKTP SSYEYNVNAT
151  RTVVNFSHAC GVSVEGEIGV LGNLETGEAG EEDGVGAVGK LSHDQMLTSV
201  EDAVRFVKDT GVDALAIAVG TSHGAYKFTR PPTGDVLRID RIKEIHQALP
251  NTHIVMHGSS SVPQEWLKVI NEYGGNIGET YGVPVEEIVE GIKHGVRKVN
301  IDTDLRLAST GAVRRYLAEN PSDFDPRKYL SKTIEAMKQI CLDRYLAFGC
351  EGQAGKIKPV SLEKMANRYA KGELNQIVK*
```

m034/a034 96.9% identity over a 257 aa overlap

```
                    10        20        30        40        50        60
m034.pep    MSCLWFFAVKNIIIRLIYLLPKETQMALVSMRQLLDHAAEXSYGLPAFNVNNLXQMRAIM
            || |||||:|||||||||||||||||||||||||||||||| ||||||||||| ||||||
a034        MSRLWFFAAKNIIIRLIYLLPKETQMALVSMRQLLDHAAENSYGLPAFNVNNLEQMRAIM
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m034.pep    EAADQVDAPVIVQASAGARKYAGAPFLRHLILAAVEVFPHIPVVMHQDHGASPDVCQRSI
            ||||||:||||||||||||||||||||||||||||| |||||||||||||||||||||||
a034        EAADQVNAPVIVQASAGARKYAGAPFLRHLILAAVEEFPHIPVVMHQDHGASPDVCQRSI
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m034.pep    QLGFSSVMMDGSLMEDGKTPSSYEYNVNATRTVVNFSHACGVSVEGEIGVLGNLETGDAG
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||
a034        QLGFSSVMMDGSLMEDGKTPSSYEYNVNATRTVVNFSHACGVSVEGEIGVLGNLETGEAG
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m034.pep    EEDGVGAVGKLSHDQMLTSVEDAVCFVKDTGVDALAIAVGTSHGAYKFTRPPTGDVLRID
            |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
a034        EEDGVGAVGKLSHDQMLTSVEDAVRFVKDTGVDALAIAVGTSHGAYKFTRPPTGDVLRID
                   190       200       210       220       230       240

                   250
m034.pep    RIKEIHQALPNTHIVMH
            |||||||||||||||||
a034        RIKEIHQALPNTHIVMHGSSSVPQEWLKVINEYGGNIGETYGVPVEEIVEGIKHGVRKVN
                   250       260       270       280       290       300
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 034 shows 96.5% identity over a 257 aa overlap with a predicted ORF (ORF 034.ng) from *N. gonorrhoeae:*

m034/g034

```
m034.pep    MSCLWFFAVKNIIIRLIYLLPKETQMALVSMRQLLDHAAEXSYGLPAFNVNNLXQMRAIM   60
            || ||||||||||||||||||||||||||||||||||||| |||||||||||| ||||||
g034        MSRLWFFAVKNIIIRLIYLLPKETQMALVSMRQLLDHAAENSYGLPAFNVNNLEQMRAIM   60

m034.pep    EAADQVDAPVIVQASAGARKYAGAPFLRHLILAAVEVFPHIPVVMHQDHGASPDVCQRSI  120
            ||||||:||||||||||||||||||||||||||||| |||||||||||||||||||||||
g034        EAADQVNAPVIVQASAGARKYAGAPFLRHLILAAVEEFPHIPVVMHQDHGASPDVCQRSI  120

m034.pep    QLGFSSVMMDGSLMEDGKTPSSYEYNVNATRTVVNFSHACGVSVEGEIGVLGNLETGDAG  180
            |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||:||
g034        QLGFSSVMMDGSLLEDGKTPSSYEYNVNATRTVVNFSHACGVSVEGEIGVLGNLETGEAG  180

m034.pep    EEDGVGAVGKLSHDQMLTSVEDAVCFVKDTGVDALAIAVGTSHGAYKFTRPPTGDVLRID  240
            |||||||:|||||||||||||||| |||||||||||||||||||||||||||||||||||
g034        EEDGVGAAGKLSHDQMLTSVEDAVRFVKDTGVDALAIAVGTSHGAYKFTRPPTGDVLRID  240

m034.pep    RIKEIHQALPNTHIVMH                                            257
            |||||||||||||||||
g034        RIKEIHQALPNTHIVMHGSSSVPQEWLKVINEYGGNIGETYGVPVEEIVEGIKHGVRKVN  300
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 127>:

g036.seq

```
  1  ATGCTGAAGC CGTGTTTGGT ATACAGTGCC TGTGCGGCGG cgttgcCTGC
 51  GCGGACTTCG AGCAGCAGGC GTTGCGTGCC TTCGGGCAGA TGTGCGTACC
101  AATATTCGAG CAGGGCGGAC GCAACGCCCC GTCGGCGGCA TTCGGGCGCG
151  GTGGCAATCA GGTGCAGTTC GGATTCGTCG GGCAGGTTCT GCCAAACGAT
201  AAAGGCGGCA ATCCTGCCGT CTTTTTCCGC AAGGAAAACC TGTTCGGACG
251  GCGAAACAAG CGCGGACTCA AATTGGCGTT GCGTCCACGC GGACGGGTTG
301  CAGACGGTAT CGAGCGCGGC CAGTGCGGCG CAGTCGGACG GTGAGGCTGG
351  GCGGATGTTC ATGTTCGTGC CTTCCGTTCC GCCTGTTCTT TGGCAGTCAG
401  GGCGATTTTG TTGCGGACGT AGAGCAGTTC GGCGTGTGCC GCGCCAGTTG
451  CGGGATAGCC GCCGCCGAGG GCGAGCGCGA GAAAATCGGC GGCGGTCGGC
501  ATATCGGGTT TGCCTGAGAA GGGCGGACGG TTTTCCAGTG CGAACGCACT
551  GCCGATGCCG TCTGAAAAGA CGTACCCCTC GGGGAGGGCA ATGTCTGCCG
601  CCCTACCGAC TTGATAATCG CTCAAACGGC GGCGGTTCAG CGTGTCGAAC
651  CACGCATAAA ACACTTCGCC CATACGCGCG TCCGCAGCGG CGAGTATGCA
701  GCTTTGCGGC GGCGGCAGCG AGGCGGCGGC ATCGAGCGTG GGGATGCCGA
751  TTAAAGGCGT GTCGAACGGC GTTGCCAAAC CTTGCGCCAC GCCGATGCCG
801  ATACGCAGTC CGGTAA
```

This corresponds to the amino acid sequence <SEQ ID 128; ORF 036.ng>:

g036.pep

```
  1  MLKPCLVYSA CAAALPARTS SSRRCVPSGR CAYQYSSRAD ATPRRRHSGA
 51  VAIRCSSDSS GRFCQTIKAA ILPSFSARKT CSDGETSADS NWRCVHADGL
101  QTVSSAASAA QSDGEAGRMF MFVPSVPPVL WQSGRFCCGR RAVRRVPRQL
151  RDSRRRGRAR ENRRRSAYRV CLRRADGFPV RTHCRCRLKR RTPRGGQCLP
201  PYRLDNRSNG GGSACRTTHK TLRPYARPQR RVCSFAAAAA RRRHRAWGCR
251  LKACRTALPN LAPRRCRYAV R*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 129>:

m036.seq

```
  1  ATGCTGAAGC CGTGCGCCGT GTACAGTGCC TGTGCGGCGG TGTTGCCTGC
 51  ACGGACTTCG AGCAGCAGGC GTTGCGTGTC TTCGGGCAGA TGTGTGAACC
101  AATATTCGAG CAGGGCGGAC GCAATTCCTT GGCGGCGGCA TTCGGGCGCG
151  GTGGCAATCA GGTGCAGTTC GGATTCGTCG GGCAGGTTCT GCCAAACGAT
201  AAAGGCGGCA ATCCCg.CGT CTTTTTCCGC AAGGAAAACC TGTTCGGACG
251  GCGAAACCAG TGCGGACTCA AATTGGCGTT GCGTCCATGC GGACGGGTTG
301  CAGACGGCAT CGAGTGCGGC CAGCTCCTCA CAATCGGCAC AAACGGCACG
```

```
351  GCGGATGTTC ACGGGCGCGC TCTCCGTTCG GCCTGTTCTT TGGCAGTCAG
401  GGCGATTTTG TTGCGGACGT AGAGCAAACC GGCGTGTGCG GCATGGACGG
451  CAGGATAACC GCCCTTGGCT GCCAATGCGA GAAAGTCGGC GGCAGTCGGC
501  ATATCCGGTC TGCCTGAGAA CGGCGGAGCT TCTTCCAGCG CGAACGCGCT
551  GCCTATGCCG TCTGAAAAGG CGCATCCCTC CGGCAGCCGG ATGTCTGCCG
601  CCCGCCCGAC CTGATAATCG CTCAAACGGT GGCAGTTCAG CGTATCGAAC
651  CATGCATAAA ACACTTCGCC CATACGAGCG TCCGTAGCGG CAAGGATGCA
701  GCTTTGCGGC GGCGGCAGCG AGGCGGCGGC ATCGAGCGAG GGTACGCCGA
751  TTAAGGGGGT ATCAAACGGC GTTGCCAAAC CCTGAGCTAC ACCGATGCCG
801  ATACGCAGTC CGGTAA
```

This corresponds to the amino acid sequence <SEQ ID 130; ORF 036>:

m036.pep

```
  1  MLKPCAVYSA CAAVLPARTS SSRRCVSSGR CVNQYSSRAD AIPWRRHSGA
 51  VAIRCSSDSS GRFCQTIKAA IPXSFSARKT CSDGETSADS NWRCVHADGL
101  QTASSAASSS QSAQTARRMF TGALSVRPVL WQSGRFCCGR RANRRVRHGR
151  QDNRPWLPMR ESRRQSAYPV CLRTAELLPA RTRCLCRLKR RIPPAAGCLP
201  PARPDNRSNG GSSAYRTMHK TLRPYERP*R QGCSFAAAAA RRRHRARVRR
251  LRGYQTALPN PELHRCRYAV R*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 131>:

a036.seq

```
  1  ATGCTGAAGC CGTGCGCCGT GTACAGTGCC TGTGCGGCGG TGTTGCCTGC
 51  ACGGACTTCG AGCAGCAGGC GTTGCGTGTC TTCGGGCAGA TGTGTGAACC
101  AATATTCGAG CAGGGCGGAC GCAATTCCTT GGCGGCGGCA TTCGGGCGCG
151  GTGGCAATCA GGTGCAGTTC GGATTCGTCG GGCAGGTTCT GCCAAACGAT
201  AAAGGCGGCA ATCCCGCCGT CTTTTTCCGC AAGGAAAACC TGTTCGGACG
251  GCGAAACCAG TGCGGACTCA AATTGGCGTT GCGTCCACGC GGACGGGTTG
301  CAGACGGCAT CGAGCGCGGC GAGTGCGGCG CAATCGGCAT AAACGGCGCG
351  GCGGATGTTC ACAGGCGCGC CCTCCGTTCC GCCTGTTCTT TGGCAGTCAA
401  GGCGATTTTG TTGCGGACGT AGAGCAGCTC GGCGTGTGCC GCAGCGACGG
451  CGGGAAAACC GCCTTCAGCC GCCAGATTGA GGAAGTCGGC GGCGGTCGGC
501  ATATCGGGTT TGCCTGAGAA GGGCGGACGG TTTTCCAGCG CGAACGCATT
551  GCCGATGCCG TCTGAAAAGG CGCATCCTTC CGGCAGCCGG ATGTCTGCCG
601  CCCGACCGAC CTGATAATCG CTCAAACGGC GGCGGTTCAG CGTGTCGAAC
651  CATGCATAAA ACACTTCGCC CATACGTGCG TCCGCAGCGG CAAGGATGCA
701  GCTTTGCGGC GGCGGCAGCG AGGCGGCGGC ATCGAGCGAG GGTACGCCGA
751  TTAAAGGAGT ATCAAACGGC GTTGCCAAAC CTTGCGCCAC GCCGATGCCG
801  ATACGCAGTC CCGTAA
```

This corresponds to the amino acid sequence <SEQ ID 132; ORF 036.a>:

a036.pep

```
  1  MLKPCAVYSA CAAVLPARTS SSRRCVSSGR CVNQYSSRAD AIPWRRHSGA
 51  VAIRCSSDSS GRFCQTIKAA IPPSFSARKT CSDGETSADS NWRCVHADGL
101  QTASSAASAA QSA*TARRMF TGAPSVPPVL WQSRRFCCGR RAARRVPQRR
151  RENRLQPPD* GSRRRSAYRV CLRRADGFPA RTHCRCRLKR RILPAAGCLP
201  PDRPDNRSNG GGSACRTMHK TLRPYVRPQR QGCSFAAAAA RRRHRARVRR
251  LKEYQTALPN LAPRRCRYAV P*
```

m036/a036 85.6% identity over a 270 aa overlap

```
                    10        20        30        40        50        60
m036.pep    MLKPCAVYSACAAVLPARTSSSRRCVSSGRCVNQYSSRADAIPWRRHSGAVAIRCSSDSS
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a036        MLKPCAVYSACAAVLPARTSSSRRCVSSGRCVNQYSSRADAIPWRRHSGAVAIRCSSDSS
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m036.pep    GRFCQTIKAAIPXSFSARKTCSDGETSADSNWRCVHADGLQTASSAASSSQSAQTARRMF
            |||||||||||| |||||||||||||||||||||||||||||||||||::||| ||||||
a036        GRFCQTIKAAIPPSFSARKTCSDGETSADSNWRCVHADGLQTASSAASAAQSAXTARRMF
                    70        80        90       100       110       120
```

```
                   130       140       150       160       170       180
m036.pep    TGALSVRPVLWQSGRFCCGRRANRRVRHGRQDNRPWLPMRESRRQSAYPVCLRTAELLPA
            ||| || |||||| |||||||| ||| : |::||   |   |||:||| |||| |: :||
a036        TGAPSVPPVLWQSRRFCCGRRAARRVPQRRRENRLQPPDXGSRRRSAYRVCLRRADGFPA
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m036.pep    RTRCLCRLKRRIPPAAGCLPPARPDNRSNGGSSAYRTMHKTLRPYERPXRQGCSFAAAAA
            ||:| ||||||| |||||||| |||||||||:|| |||||||||| || |||||||||||
a036        RTHCRCRLKRRILPAAGCLPPDRPDNRSNGGGSACRTMHKTLRPYVRPQRQGCSFAAAAA
                   190       200       210       220       230       240

                   250       260       270
m036.pep    RRRHRARVRRLRGYQTALPNPELHRCRYAVRX
            |||||||||||: |||||||   :||||||
a036        RRRHRARVRRLKEYQTALPNLAPRRCRYAVPX
                   250       260       270
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

ORF 036 shows 74.9% identity over a 271 aa overlap with a predicted ORF (ORF 036.ng) from *N. gonorrhoeae:*

m036/g036

```
                    10        20        30        40        50        60
m036.pep    MLKPCAVYSACAAVLPARTSSSRRCVSSGRCVNQYSSRADAIPWRRHSGAVAIRCSSDSS
            ||||| |||||||:|||||||||||| ||||: |||||||| | ||||||||||||||||
g036        MLKPCLVYSACAAALPARTSSSRRCVPSGRCAYQYSSRADATPRRRHSGAVAIRCSSDSS
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m036.pep    GRFCQTIKAAIPXSFSARKTCSDGETSADSNWRCVHADGLQTASSAASSSQSAQTARRMF
            |||||||||||  |||||||||||||||||||||||||||||:|||||::||   | |||
g036        GRFCQTIKAAILPSFSARKTCSDGETSADSNWRCVHADGLQTVSSAASAAQSDGEAGRMF
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m036.pep    TGALSVRPVLWQSGRFCCGRRANRRVRHGRQDNRPWLPMRESRRQSAYPVCLRTAELLPA
              : || ||||||||||||||| ||| :  :|:|      |:||:||| |||| |: :|:
g036        MFVPSVPPVLWQSGRFCCGRRAVRRVPRQLRDSRRRGRARENRRRSAYRVCLRRADGFPV
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m036.pep    RTRCLCRLKRRIPPAAGCLPPARPDNRSNGGSSAYRTMHKTLRPYERPXRQGCSFAAAAA
            ||:| |||||| | :: |||| | |||||||:|| || ||||||| || |: ||||||||
g036        RTHCRCRLKRRTPRGGQCLPPYRLDNRSNGGGSACRTTHKTLRPYARPQRRVCSFAAAAA
                   190       200       210       220       230       240

                   250       260       270
m036.pep    RRRHRARVRRLRGYQTALPNPELHRCRYAVRX
            ||||||   ||:: :|||||   :|||||||
g036        RRRHRAWGCRLKACRTALPNLAPRRCRYAVRX
                   250       260       270
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 133>:

```
m036-1.seq
       1  ATGCTGAAGC CGTGCGCCGT GTACAGTGCC TGTGCGGCGG TGTTGCCTGC
      51  ACGGACTTCG AGCAGCAGGC GTT^^GTGTC TTCGGGCAGA TGTGTGAACC
     101  AATATTCGAG CAGGGCGGAC GCAATTCCTT GGCGGCGGCA TTCGGGCGCG
     151  GTGGCAATCA GGTGCAGTTC GGATTCGTCG GGCAGGTTCT GCCAAACGAT
     201  AAAGGCGGCA ATCCCGCCGT CTTTTTCCGC AAGGAAAACC TGTTCGGACG
     251  GCGAAACCAG TGCGGACTCA AATTGGCGTT GCGTCCATGC GGACGGGTTG
     301  CAGACGGCAT CGAGTGCGGC CAGCTCCTCA CAATCGGCAC AAACGGCACG
```

```
351   GCGGATGTTC ACGGGCGCGC TCTCCGTTCG GCCTGTTCTT TGGCAGTCAG
401   GGCGATTTTG TTGCGGACGT AGAGCAAACC GGCGTGTGCG GCATGGACGG
451   CAGGATAACC GCCCTTGGCT GCCAATGCGA GAAAGTCGGC GGCAGTCGGC
501   ATATCCGGTC TGCCTGAGAA CGGCGGAGCT TCTTCCAGCG CGAACGCGCT
551   GCCTATGCCG TCTGAAAAGG CGCATCCCTC CGGCAGCCGG ATGTCTGCCG
601   CCCGC[illegible]CGAC CTGATAATCG CTCAAACGGT GGCAGTTCAG CGTATCGAAC
651   CATGCATAAA ACACTTCGCC CATACGAGCG TCCGTAG
```

This corresponds to the amino acid sequence <SEQ ID 134; ORF 0036-1>:

```
m036-1.pep
    1   MLKPCAVYSA CAAVLPARTS SSRRCVSSGR CVNQYSSRAD AIPWRRHSGA
   51   VAIRCSSDSS GRFCQTIKAA IPPSFSARKT CSDGETSADS NWRCVHADGL
  101   QTASSAASSS QSAQTARRMF TGALSVRPVL WQSGRFCCGR RANRRVRHGR
  151   QDNRPWLPMR ESRRQSAYPV CLRTAELLPA RTRCLCRLKR RIPPAAGCLP
  201   PARPDNRSNG GSSAYRTMHK TLRPYERP*
```

m036-1/g036 76.8% identity in 228 aa overlap

```
                       10        20        30        40        50        60
m036-1.pep    MLKPCAVYSACAAVLPARTSSSRRCVSSGRCVNQYSSRADAIPWRRHSGAVAIRCSSDSS
              ||||| |||||||:|||||||||||| ||||: |||||||| | ||||||||||||||||
g036          MLKPCLVYSACAAALPARTSSSRRCVPSGRCAYQYSSRADATPRRRHSGAVAIRCSSDSS
                       10        20        30        40        50        60

                       70        80        90       100       110       120
m036-1.pep    GRFCQTIKAAIPPSFSARKTCSDGETSADSNWRCVHADGLQTASSAASSSQSAQTARRMF
              ||||||||||| ||||||||||||||||||||||||||||||:|||||::||   | |||
g036          GRFCQTIKAAILPSFSARKTCSDGETSADSNWRCVHADGLQTVSSAASAAQSDGEAGRMF
                       70        80        90       100       110       120

                      130       140       150       160       170       180
m036-1.pep    TGALSVRPVLWQSGRFCCGRRANRRVRHGRQDNRPWLPMRESRRQSAYPVCLRTAELLPA
               :  || ||||||||||||||| ||| :  :|:|     ||:||:||| |||| |: :|:
g036          MFVPSVPPVLWQSGRFCCGRRAVRRVPRQLRDSRRRGRARENRRRSAYRVCLRRADGFPV
                      130       140       150       160       170       180

                      190       200       210       220      229
m036-1.pep    RTRCLCRLKRRIPPAAGCLPPARPDNRSNGGSSAYRTMHKTLRPYERPX
              ||:| |||||| | :: |||| | |||||||:|| || ||||||| ||
g036          RTHCRCRLKRRTPRGGQCLPPYRLDNRSNGGGSACRTTHKTLRPYARPQRRVCSFAAAAA
                      190       200       210       220       230       240

g036          RRRHRAWGCRLKACRTALPNLAPRRCRYAVRX
                      250       260       270
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 135>:

```
g038.seq
    1   ATGACTGATT TCCGCCAAGA TTTCCTCAAA TTCTCCCTCG CCCAAAATGT
   51   TTTGAAATTC GGCGAATTTA CCACCAAAGC CGGACGGCGG TCGCCCTATT
  101   TCTTCAATGC CGGCCTCTTC AACGACGGCG CGTCCACGCT GCAACTGGCA
  151   AAATTCTATG CACAATCCAT CATTGAAAGC GGCATCCGAT TCGATATGCT
  201   GTTCGGCCCC GCCTACAAAG GCATTATTTT GGCGGCGGCA ACCGCGATGA
  251   TGCTGGCGGA AAAAGGCGTG AACGTCCCGT TTGCCTACAA CCGCAAAGAA
  301   GCCAAAGACC GCGGCGAAGG CGGCGTGTTG GTCGGCGCGC CGCTTAAAGG
  351   GCGCGTGCTG ATTATCGACG ACGTGATTTC CGCCGGCACA TCCGTACGCG
  401   AATCAATCAA ACTGATTGAA GCGGAGGGTG CAACCCCCGC CGGTGTCGCC
  451   ATCGCGCTCG ACCGCATGGA AAAAGGCACG GGTAAATTGT CCGCCGTTCA
  501   GGAAGTGGAA AAACAATACG GCCTGCCCGT CGCCCCCATC GCCAGCCTGA
  551   ACGATTTGTT TATCCTGTTG CAAAACAACC CCGAATTCGG ACAGTTCCTC
  601   GAACCCGTCC GCACCTACCG CCGGCAGTAC GGCGTAGAAT AA
```

This corresponds to the amino acid sequence <SEQ ID 136; ORF 038.ng>:

```
g038.pep
     1  MTDFRQDFLK FSLAQNVLKF GEFTTKAGRR SPYFFNAGLF NDGASTLQLA
    51  KFYAQSIIES GIRFDMLFGP AYKGIILAAA TAMMLAEKGV NVPFAYNRKE
   101  AKDRGEGGVL VGAPLKGRVL IIDDVISAGT SVRESIKLIE AEGATPAGVA
   151  IALDRMEKGT GKLSAVQEVE KQYGLPVAPI ASLNDLFILL QNNPEFGQFL
   201  EPVRTYRRQY GVE*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 137>:

```
m038.seq
     1  ATGACCGATT TCCGCCAAGA TTTCCTCAAA TTCTCCCTCG CCCAAAATGT
    51  TTTGAAATTC GGCGAATTTA CCACCAAGGC AGGACGGCGG TCGCCCTATT
   101  TCTTCAATGC CGGCCTCTTT AACGACGGCT TGTCCACGCT GCAACTGGCA
   151  AAATTTTACG CACAATCCAT CATTGAAAGC GGCATCCGAT TCGATATGCT
   201  GTTCGGTCCC GCCTACAAAG GCATTATTTT GGCGGCGGCA ACCGCGATGA
   251  TGCTGGCGGA AAAAGGCGTG AACGTCCCGT TTGCCTACAA CCGCAAAGAA
   301  GCCAAAGACC ACGGCGAAGG CGGCGTGTTG GTCGGCGCGC CGCTTAAAGG
   351  GCGCGTGCTG ATTATCGACG ACGTGATTTC CGCCGGCACA TCCGTACGCG
   401  AATCGATCAA ACTGATTGAA GCGGAGGGTG CAACCCCcGC CGGTGTCGCC
   451  ATCGCGCTCG ATCGCATGGA AAAAGGCACG GGTGAATTGA GCGCGGTTCA
   501  GGAAGTGGAr AAACAATACG GkCTGCCCGT CGCCCCCATC GCCAGCCTGA
   551  ACGATTTGTT TATTCTGTTG CAAAACAACC CCGAATTCGG ACAGTTCCTC
   601  GAACCCGTCC GAGCCTACCG TCGGCAGTAC GGCGTAGAAT AA
```

This corresponds to the amino acid sequence <SEQ ID 138; ORF 038>:

```
m038.pep
     1  MTDFRQDFLK FSLAQNVLKF GEFTTKAGRR SPYFFNAGLF NDGLSTLQLA
    51  KFYAQSIIES GIRFDMLFGP AYKGIILAAA TAMMLAEKGV NVPFAYNRKE
   101  AKDHGEGGVL VGAPLKGRVL IIDDVISAGT SVRESIKLIE AEGATPAGVA
   151  IALDRMEKGT GELSAVQEVE KQYGLPVAPI ASLNDLFILL QNNPEFGQFL
   201  EPVRAYRRQY GVE*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 139>:

```
a038.seq
     1  ATGACCGATT TCCGCCAAGA TTTCCTCAAA TTCTCCCTCG CCCAAAATGT
    51  TTTGAAATTC GGCGAATTCA CCACCAAAGC CGGACGGCGG TCGCCCTATT
   101  TCTTCAATGC CGGCCTCTTT AACGACGGCT TGTCCACGCT GCAACTGGCA
   151  AAATTTTACG CACAATCCAT CATTGAAAGC GGCATCCGAT TCGATATGCT
   201  GTTCGGCCCC GCCTACAAAG GCATTATTTT GGCGGCGGCA ACCGCGATGA
   251  TGCTGGCGGA AAAAGGCGTG AACGTCCCCGT TTGCCTACAA CCGCAAAGAA
   301  GCCAAAGACC ACGGCGAAGG CGGCGTGTTG GTCGGCGCGC CGCTTAAAGG
   351  GCGCGTGCTG ATTATCGACG ACGTGATTTC CGCCGGCACA TCCGTACGCG
   401  AATCGATCAA ACTGATTGAA GCGGAGGGTG CAACCCCCGC CGGTGTCGCC
   451  ATCGCGCTCG ACCGCATGGA AAAAGGCACG GGTGAATTGA GCGCGGTTCA
   501  GGAAGTGGAA AAACAATACG GCCTGCCCGT CGCCCCCATC GCCAGCCTGA
   551  ACGATTTGTT TATTCTGTTG CAAAACAACC CCGAATTCGG ACAGTTCCTC
   601  GAACCCGTCC GAGCCTACCG TCGGCAGTAC GGCGTAGAAT AA
```

This corresponds to the amino acid sequence <SEQ ID 140; ORF 038.a>:

```
a038.pep
     1  MTDFRQDFLK FSLAQNVLKF GEFTTKAGRR SPYFFNAGLF NDGLSTLQLA
    51  KFYAQSIIES GIRFDMLFGP AYKGIILAAA TAMMLAEKGV NVPFAYNRKE
   101  AKDHGEGGVL VGAPLKGRVL IIDDVISAGT SVRESIKLIE AEGATPAGVA
   151  IALDRMEKGT GELSAVQEVE KQYGLPVAPI ASLNDLFILL QNNPEFGQFL
   201  EPVRAYRRQY GVE*
```

m038/a038 100.0% identity over a 213 aa overlap

```
                    10        20        30        40        50        60
m038.pep    MTDFRQDFLKFSLAQNVLKFGEFTTKAGRRSPYFFNAGLFNDGLSTLQLAKFYAQSIIES
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a038           MTDFRQDFLKFSLAQNVLKFGEFTTKAGRRSPYFFNAGLFNDGLSTLQLAKFYAQSIIES
                      10        20        30        40        50        60

                      70        80        90       100       110       120
m038.pep       GIRFDMLFGPAYKGIILAAATAMMLAEKGVNVPFAYNRKEAKDHGEGGVLVGAPLKGRVL
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a038           GIRFDMLFGPAYKGIILAAATAMMLAEKGVNVPFAYNRKEAKDHGEGGVLVGAPLKGRVL
                      70        80        90       100       110       120

                     130       140       150       160       170       180
m038.pep       IIDDVISAGTSVRESIKLIEAEGATPAGVAIALDRMEKGTGELSAVQEVEKQYGLPVAPI
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a038           IIDDVISAGTSVRESIKLIEAEGATPAGVAIALDRMEKGTGELSAVQEVEKQYGLPVAPI
                     130       140       150       160       170       180

                     190       200       210
m038.pep       ASLNDLFILLQNNPEFGQFLEPVRAYRRQYGVEX
               ||||||||||||||||||||||||||||||||||
a038           ASLNDLFILLQNNPEFGQFLEPVRAYRRQYGVEX
                     190       200       210
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 038 shows 98.1% identity over a 213 aa overlap with a predicted ORF (ORF 038.ng) from *N. gonorrhoeae:*

m038/g038

```
                      10        20        30        40        50        60
m038.pep       MTDFRQDFLKFSLAQNVLKFGEFTTKAGRRSPYFFNAGLFNDGLSTLQLAKFYAQSIIES
               |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
g038           MTDFRQDFLKFSLAQNVLKFGEFTTKAGRRSPYFFNAGLFNDGASTLQLAKFYAQSIIES
                      10        20        30        40        50        60

                      70        80        90       100       110       120
m038.pep       GIRFDMLFGPAYKGIILAAATAMMLAEKGVNVPFAYNRKEAKDHGEGGVLVGAPLKGRVL
               |||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
g038           GIRFDMLFGPAYKGIILAAATAMMLAEKGVNVPFAYNRKEAKDRGEGGVLVGAPLKGRVL
                      70        80        90       100       110       120

                     130       140       150       160       170       180
m038.pep       IIDDVISAGTSVRESIKLIEAEGATPAGVAIALDRMEKGTGELSAVQEVEKQYGLPVAPI
               |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
g038           IIDDVISAGTSVRESIKLIEAEGATPAGVAIALDRMEKGTGKLSAVQEVEKQYGLPVAPI
                     130       140       150       160       170       180

                     190       200       210
m038.pep       ASLNDLFILLQNNPEFGQFLEPVRAYRRQYGVEX
               ||||||||||||||||||||||||:|||||||||
g038           ASLNDLFILLQNNPEFGQFLEPVRTYRRQYGVEX
                     190       200       210
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 141>:

g039.seq

```
  1  ATGCCGTCCG AACCACCTGC CGCTTCAGAC GGCATCAAAC CGACACACAC
 51  CGAGAAAACA TCATGCCCGC CTGTTTCTGT CCGCACTGCA AAACCCGCCT
101  CTGGGTCAAA GAAAcccagC TCAAcgtCgC ccaagGCTTC GTCGTCTgcc
151  aaAAAtgcga agGGCTgttt aaAgccaaaG accAtctggc aaGcacGAAA
201  gaacctatat tcaacgattg gcccgaagct gtttcgggat gTcaaaCTCG
```

```
251   TCcaccgcaT cggcacgcac gccattagca aGAaacagat gtcccgcgac
301   gaaatCgccg atatcctcaa cggcggtaca acCCTGCACG ATACGCCGCC
351   CGCAACCGCC GCTGCCGCac ctGCCGCCGC ACCGCaggTT TCCGTACCGC
401   CCGCCCGTCA GGAAGGGCTC AACTGGACTA TTGCAACCCT GTTCGCACTT
451   ATCGTCCTCA TTATGCAGCT TTCCTACCTC TTCATCCTAT GA
```

This corresponds to the amino acid sequence <SEQ ID 142; ORF 039.ng>:

g039.pep

```
  1   MPSEPPAASD GIKPTHTEKT SCPPVSVRTA KPASGSKKPS STSPKASSSA
 51   KNAKGCLKPK TIWQARKNLY STIGPKLFRD VKLVHRIGTH AISKKQMSRD
101   EIADILNGGT TLHPTPPATA AAAPAAAPQV SVPPARQEGL NWTIATLFAL
151   IVLIMQLSYL FIL*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 143>:

m039.seq

```
  1   ATGCCGTCCG AACCGCCTTA CGCCTCAGAC GGCATCAAAC CTGACACACA
 51   CGAGGAAATA CCATGCCCGC CTGTTTCTGC CCCCACTGCA AAACCCGTCT
101   CTGGGTCAAA GAAACCCAAC TCAATGTCGC CGnnnnnnnn nnnnnnnnnn
151   nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
201   nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnCCC GAGGCTGTTT
251   CGGATGTCAA ACTCGTTCAC CGTATCGGCA CGCGCGCCAT CGGCAAGAAA
301   CAGATTTCCC GTGACGAAAT CGCCGGCATC CTCAACGGCG GTACAACCCA
351   GCCCGATATT CCGCCCGCAA CCGCCGCCAC CCCTGCTGCC GCACCGCAGG
401   TTACCGTACC GCCCGCCGCG CCCGCCCGTC AGGATGGGTT CAACTGGACG
451   ATTGCAACCC TGTTTGCCCT TATCGTCCTC ATTATGCAGC TTTCCTACCT
501   CGTCATCCTA TGA
```

This corresponds to the amino acid sequence <SEQ ID 144; ORF 039>:

m039.pep

```
  1   MPSEPPYASD GIKPDTHEEI PCPPVSAPTA KPVSGSKKPN SMSPXXXXXX
 51   XXXXXXXXXX XXXXXXXXXX XXXXXXXXXP EAVSDVKLVH RIGTRAIGKK
101   QISRDEIAGI LNGGTTQPDI PPATAATPAA APQVTVPPAA PARQDGFNWT
151   IATLFALIVL IMQLSYLVIL *
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 145>:

a039.seq

```
  1   ATGCCGTCTG AACCGCCTTA CGCCTCAGAC GGCATCAAAC CTGACACACA
 51   CGAGGAAATA CCATGCCCGC CTGTTTCTGC CCCCACTGCA AAACCCGTCT
101   CTGGGTCAAA GAAACCCAAC TCAATGTCGC CCAAGGCTTC GTCGTCTGCC
151   AAAAATGCGA AGGAATGTTT AAAGCCAAAG ACCATCTGGC AAGCACGAAA
201   GAACCCATAT TCAACGATT. TGCCCGAAGC TGTTTCGGAT GTCAAACTCG
251   TTCACCGCAT CGGCACGAGC GCCATCGGCA AGAAACAGAT TTCCCGTGAC
301   GAAATCGCCG GCATCCTCAA CGGCGGCACA ACCCAGCCCG ATATTCCGCC
351   CGCAACCGCC GCCACCCCTG CTGCCGCACC GCAGGTTACC GTACCGCCCG
401   CCGCGCCCGC CCGTCAGGAT GGGTTCAACT GGACGATTGC AACCCTGTTT
451   GCCCTTATCG TCCTCATTAT GCAGCTTTCC TACCTCGTCA TCCTATGA
```

This corresponds to the amino acid sequence <SEQ ID 146; ORF 039.a>:

a039.pep

```
  1   MPSEPPYASD GIKPDTHEEI PCPPVSAPTA KPVSGSKKPN SMSPKASSSA
 51   KNAKECLKPK TIWQARKNPY STIXPEAVSD VKLVHRIGTS AIGKKQISRD
101   EIAGILNGGT TQPDIPPATA ATPAAAPQVT VPPAAPARQD GFNWTIATLF
151   ALIVLIMQLS YLVIL*
```

m039/a039 79.4% identity over a 170 aa overlap

```
                        10        20        30        40        50        60
m039.pep     MPSEPPYASDGIKPDTHEEIPCPPVSAPTAKPVSGSKKPNSMSPXXXXXXXXXXXXXXXX
             ||||||||||||||||||||||||||||||||||||||||||||
a039         MPSEPPYASDGIKPDTHEEIPCPPVSAPTAKPVSGSKKPNSMSPKASSSAKNAKECLKPK
```

```
                      10        20        30        40        50        60

                      70        80        90       100       110       120
m039.pep      XXXXXXXXXXXXXXXXXXPEAVSDVKLVHRIGTRAIGKKQISRDEIAGILNGGTTQPDI
              :           :|    ||||||||||||||| ||||||||||||||||||||||||||
a039          TIWQARKNPYSTIX-----PEAVSDVKLVHRIGTSAIGKKQISRDEIAGILNGGTTQPDI
                      70            80        90       100       110

                     130       140       150       160       170
m039.pep      PPATAATPAAAPQVTVPPAAPARQDGFNWTIATLFALIVLIMQLSYLVILX
              |||||||||||||||||||||||||||||||||||||||||||||||||||
a039          PPATAATPAAAPQVTVPPAAPARQDGFNWTIATLFALIVLIMQLSYLVILX
               120       130       140       150       160
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 039 shows 60.8% identity over a 171 aa overlap with a predicted ORF (ORF 039.ng) from *N. gonorrhoeae:*

m039/g039

```
                      10        20        30        40        50        60
m039.pep      MPSEPPYASDGIKPDTHEEIPCPPVSAPTAKPVSGSKKPNSMSPXXXXXXXXXXXXXXXX
              |||||| |||||||   |:  |||||: ||||:||||||:| ||
g039          MPSEPPAASDGIKPTHTEKTSCPPVSVRTAKPASGSKKPSSTSPKASSSAKNAKGCLKPK
                      10        20        30        40        50        60

                      70        80        90       100       110       120
m039.pep      XXXXXXXXXXXXXXXXXXPEAVSDVKLVHRIGTRAIGKKQISRDEIAGILNGGTTQPDI
              :          :      |:   ||||||||||:||:|||:|||||| |||||||  |
g039          TIWQARKNLYSTIG-----PKLFRDVKLVHRIGTHAISKKQMSRDEIADILNGGTTLHDT
                      70            80        90       100       110

                     130       140       150       160       170
m039.pep      PPATAAT-PAAAPQVTVPPAAPARQDGFNWTIATLFALIVLIMQLSYLVILX
              ||||||: |||||||:||||   ||:|:|||||||||||||||||||| |||
g039          PPATAAAAPAAAPQVSVPPA---RQEGLNWTIATLFALIVLIMQLSYLFILX
                120       130          140       150       160
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 147>:

g040.seq

```
  1  ATGAACGCGC CCGACAGCTT TGTCGCCCAC TTCCGCGAAG CCGCCCCCTA
 51  CATCCGCCAA ATGCGCGGCA CGACACTGGT CGCCGGCATA GAcggCCGCC
101  TGCTCGAAGG CGGCACCTTA AATAAGCTCG CCGCCGACAT CGGGCTGTTG
151  TCGCAACTGG GCATCCGACT CGTCCTCATC CACGGCGCGT ACCACTTCCT
201  CGAccgCCTC GCCGCCGCGC AAGgccGCAC GCCGCATTAT TGCCGgggtt
251  tGCGCGTTAC CGACGaAACc tcGctcgGAC AGGCGCAGCA GtttGCCGGC
301  AccgTCCGCA GCCGTTTTGA agcCGCATTG tgcggcagCG tttcaggatt
351  cgcgCGCGCG CCTTCCGTCC CGCTCGTAtc gggcaacttc ctgacCGCCC
401  GTCcgatggg cgtgattgac ggaACCGata tggaatacgc gggggttatc
451  cgcaaaaccg ACACCGCCGC CCTCCGTTTC CAACTCGACG CGGGCAATAT
501  CGTCTGGATG CCGCCGCTCG GGCATTCCTA CGGCGGCAAA ACCTTCAATC
551  TCGATATGGT GCAGGCCGCC GCTTCCGTCG CCGTCTCGCT TCAGGCCGAA
601  AAACTCGTTT ACCTGACCCT TTCAGACGGC ATTTCCCGCC CCGACGGCAC
651  GCTCGCCGAA ACCCTCTCGG CACAGGAAGC GCAATCGCTG GCGGAACACG
701  CCGCCAGCGA AACCCGACGA CTGATTTCGT CCGCCGTTGC CGCGCTCGAA
751  GGCGGCGTGC ATCGCGTCCA AATCCTCAAC GGGGCCGCCG ACGGCAGCCT
801  GCTGCAAGAA CTCTTCACCC GCAACGGCAT CGGCACGTCC ATTGCCAAAG
851  AAGCCTTCGT CTCCATCCGG CAGGCGCACA GCGGCGACAT CCCGCACATC
901  GCCGCCCTCA TCCGCCCCGCT GGAAGAACAG GGCGTCCTAT TGCACCGCAG
951  CCGCGAATAC CTCGAAAACC ACATTTCCGA ATTTTCCATC CTCGAACACG
```

```
1001  ACGGCGACCT GTACGGCTGT GCCGCACTCA AAACCTTTGC CGAAGCCGAT
1051  TGCGGCGAAA TCGCCTGCCT TGCCGTCTCG CCGCAGGCAC AGGACGGCGg
1101  ctACGGCGAA CGCCTGCTTG CCCACATTAT CGATAAGGCG CGCGGCATAG
1151  GCATAAGCAG GCTGTTCGCA CTGTCCACAA ATACCGGCGA ATGGTTTGCC
1201  GAACGCGGCT TTCAGACGGC ATCGGAAGAC GAGCTGCCCG AAACGCGGCG
1251  CAAAGACTAC CGCAGCAACG GACGAAACCC GCATATTCTG GTGCGTCGCC
1301  TGCACCGCTG A
```

This corresponds to the amino acid sequence <SEQ ID 148; ORF 040.ng>:

g040.pep

```
  1  MNAPDSFVAH FREAAPYIRQ MRGTTLVAGI DHRLLEGGTL NKLAADIGLL
 51  SQLGIRLVLI HGAYHFLDRL AAAQGRTPHY CRGLRVTDET SLGQAQQFAG
101  TVRSRFEAAL CGSVSGFARA PSVPLVSGNF LTARPMGVID GTDMEYAGVI
151  RKTDTAALRF QLDAGNIVWM PPLGHSYGGK TFNLDMVQAA ASVAVSLQAE
201  KLVYLTLSDG ISRPDGTLAE TLSAQEAQSL AEHAASETRR LISSAVAALE
251  GGVHRVQILN GAADGSLLQE LFTRNGIGTS IAKEAFVSIR QAHSGDIPHI
301  AALIRPLEEQ GVLLHRSREY LENHISEFSI LEHDGDLYGC AALKTFAEAD
351  CGEIACLAVS PQAQDGGYGE RLLAHIIDKA RGIGISRLFA LSTNTGEWFA
401  ERGFQTASED ELPETRRKDY RSNGRNPHIL VRRLHR*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 149>:

m040.seq

```
   1  ATGAGCGCGC CCGACCTCTT TGTCGCCCAC TTCCGCGAAG CCGTCCCCTA
  51  CATCCGCCAA ATGCGCGGCA AAACGCTGGT CGCCGGCATA GACGACCGCC
 101  TGCTCGAAGG TGATACCTTA AACAAGCTCG CCGCCGACAT CGGGCTGTTG
 151  TCGCAACTGG GCATCAGGCT CGTCCTCATC CACGGCGCGC GCCACTTCCT
 201  CGACCGCCAC GCCGCCGCTC AAGGCCGCAC GCCGCATTAT TGCCGGGGCT
 251  TGCGCGTTAC CGACGAAACC TCGCTCGAAC AGGCGCAgCA GTTTGCCGGC
 301  ACCGTCCGCA GCCGTTTTGA AGCCGCATTG TGCGGCAGCG TTTCCGGGTT
 351  CGCGCGCGCG CCTTCCGTCC CGCTCGTATC GGGCAACTTC CTGACCGCCC
 401  GTCCGATAGG TGTGATTGAC GGAACCGATA TGGAATACGC GGGCGTTATC
 451  CGCAAAACCG ACACCGCCGC CCTCCGTTTC CAACTCGACG CGGGCAATAT
 501  CGTCTGGCTG CCGCCGCTCG GACATTCCTA CAGCGGCAAG ACCTTCTATC
 551  TCGATATGCT TCAAACCGCC GCCTCCGCCG CCGTCTCGCT TCAGGCCGAA
 601  AAACTCGTTT ACCTGACCCT TTCAGACGGC ATTTCCCGCC CCGACGGCAC
 651  GCTCGCCGAA ACCCTCTCGG CACAGGAAGC GCAATCGCTG GCGGAACACG
 701  CCGGCGGGCA AACGCGACGG CTGATTTCGT CCGCCGAACT CTTCACCCGC
 751  AACGGCATCG GCACGTCCAT TGCCAAAGAA GCCTTCGTCT CCATCCGGCA
 801  rGCGCAywgG G.CGACATCC CGCACATCGC CGCCCTCATC CGCCCGCTGG
 851  AAGAACAGGG CATCCTGCTG CACCGCAs.c GCGAATACCT CGAAAACCAC
 901  ATTTCCGAAT TTTCCATCCT CGAACACGAC GGCAACCTGT ACGGTTGCGC
 951  CGCCCTGAAA ACCTTTGCCG AAGCCGATTG CGGCGAAATC GCCTGCCTTG
1001  CCGTCTCGCC GCag.cACAG GACGGCGGCT ACGGCGAACG CnTGCTTGCC
1051  CACATTATCG ATAAGGCGCG CGGCATAGGC ATAAGCAGGC TGTTCGCACT
1101  GTCCACAAAT ACCGGCGAAT GGTTTGCCGA ACGCGGCTTT CAGACGGCAT
1151  CGGAAGACGA GTTGCCCGAA ACGCGGCGCA AAGACTACCG CAGCAACGGA
1201  CGGAACTCGC ATATTCTGGT ACGTCGCCTG CACCGCTGA
```

This corresponds to the amino acid sequence <SEQ ID 150; ORF 040>:

m040.pep

```
  1  MSAPDLFVAH FREAVPYIRQ MRGKTLVAGI DDRLLEGDTL NKLAADIGLL
 51  SQLGIRLVLI HGARHFLDRH AAAQGRTPHY CRGLRVTDET SLEQAQQFAG
101  TVRSRFEAAL CGSVSGFARA PSVPLVSGNF LTARPIGVID GTDMEYAGVI
151  RKTDTAALRF QLDAGNIVWL PPLGHSYSGK TFYLDMLQTA ASAAVSLQAE
201  KLVYLTLSDG ISRPDGTLAE TLSAQEAQSL AEHAGGQTRR LISSAELFTR
251  NGIGTSIAKE AFVSIRQAHX XDIPHIAALI RPLEEQGILL HRXREYLENH
301  ISEFSILEHD GNLYGCAALK TFAEADCGEI ACLAVSPQXQ DGGYGERXLA
351  HIIDKARGIG ISRLFALSTN TGEWFAERGF QTASEDELPE TRRKDYRSNG
401  RNSHILVRRL HR*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 151>:

a040.seq

```
   1  ATGATCGTGC CCGACCTCTT TGTCGCCCAC TTCCGCGAAG CCGCCCCCTA
  51  CATCCGCCAA ATGCGCGGCA AAACGCTGGT CGCCGGCATA GACGACCGCC
 101  TGCTCGAAGG TGATACCTTA AACAAGTTCG CCGCCGACAT CGGGCTTTTG
 151  TCGCAACTGG GCATCAGGCT CGTCCTCATC CACGGCGCGC GCCACTTCCT
 201  CGACCGCCAC GCCGCCGCGC AAGGCCGCAC GCCGCATTAT TGCCGGGGCT
 251  TGCGCGTTAC CGACGAAACC TCGCTCGAAC AGGCGCAGCA GTTTGCCGGC
 301  ACCGTCCGCA GCCGTTTTGA AGCCGCATTG TGCGGCAGCG TTTCCGGGTT
 351  CGCGCGCGCG CCTTCCGTCC CGCTCGTATC GGGCAACTTC CTGACCGCCC
 401  GTCCGATAGG TGTGATTGAC GGAACCGATA TGGAATACGC GGGCGTTATC
 451  CGCAAAACCG ACACCGCCGC CCTCCGTTTC CAACTCGACG CGGGCAATAT
 501  CGTCTGGCTG CCGCCGCTCG GACATTCCTA CAGCGGCAAG ACCTTCCATC
 551  TCGATATGCT TCAAACCGCC GCCTCCGTCG CCGTCTCGCT TCAGGCCGAA
 601  AAACTCGTTT ACCTGACCCT TTCAGACGGC ATTTCCCGCC CCGACGGCAC
 651  GCTCGCCGTA ACCCTCTCGG CACAGGAAGC GCAATCGCTG GCGGAACACG
 701  CCGGCGGCGA AACGCGACGG CTGATTTCGT CCGCCGTTGC CGCGCTCGAA
 751  GGCGGCGTGC ATCGCGTCCA AATCCTCAAC GGAGCCGCCG ACGGCAGCCT
 801  GCTGCAAGAA CTCTTCACCC GCAACGGCAT CGGCACGTCC ATTGCCAAAG
 851  AAGCCTTCGT CTCCATCCGG CAGGCGCACA GCGGCGACAT CCCGCACATT
 901  GCCGCCCTCA TCCGCCCGCT GGAAGAACAG GGCATCCTGC TGCACCGCAG
 951  CCGCGAATAC CTCGAAAACC ACATTTCCGA ATTTTCCATC CTCGAACACG
1001  ACGGCAACCT GTACGGTTGC GCCGCCCTGA AAACCTTTGC CGAAGCCGAT
1051  TGCGGCGAAA TCGCCTGCCT TGCCGTCTCG CCGCAGGCAC AGGACGGCGG
1101  CTACGGCGAA CGCCTGCTTG CCCACATTAT CGATAAGGCG CGCGGCATAG
1151  GCATAAGCAG GCTGTTCGCA CTGTCCACAA ATACCGGCGA ATGGTTTGCC
1201  GAACGCGGCT TTCAGACGGC ATCGGAAGAC GAGTTGCCCG AAACGCGGCG
1251  CAAAGACTAC CGCAGCAACG GACGGAACTC GCATATTCTG GTGCGTCGCC
1301  TGCACCGCTG A
```

This corresponds to the amino acid sequence <SEQ ID 152; ORF 040.a>:

```
a040.pep
   1  MIVPDLFVAH FREAAPYIRQ MRGKTLVAGI DDRLLEGDTL NKFAADIGLL
  51  SQLGIRLVLI HGARHFLDRH AAAQGRTPHY CRGLRVTDET SLEQAQQFAG
 101  TVRSRFEAAL CGSVSGFARA PSVPLVSGNF LTARPIGVID GTDMEYAGVI
 151  RKTDTAALRF QLDAGNIVWL PPLGHSYSGK TFHLDMLQTA ASVAVSLQAE
 201  KLVYLTLSDG ISRPDGTLAV TLSAQEAQSL AEHAGGETRR LISSAVAALE
 251  GGVHRVQILN GAADGSLLQE LFTRNGIGTS IAKEAFVSIR QAHSGDIPHI
 301  AALIRPLEEQ GILLHRSREY LENHISEFSI LEHDGNLYGC AALKTFAEAD
 351  CGEIACLAVS PQAQDGGYGE RLLAHIIDKA RGIGISRLFA LSTNTGEWFA
 401  ERGFQTASED ELPETRRKDY RSNGRNSHIL VRRLHR*
```

m040/a040 91.5% identity in 436 aa overlap

```
                        10        20        30        40        50        60
m040.pep    MSAPDLFVAHFREAVPYIRQMRGKTLVAGIDDRLLEGDTLNKLAADIGLLSQLGIRLVLI
            | :|||||||||||:|||||||||||||||||||||||||||:|||||||||||||||||
a040        MIVPDLFVAHFREAAPYIRQMRGKTLVAGIDDRLLEGDTLNKFAADIGLLSQLGIRLVLI
                        10      · 20        30        40        50        60

                        70        80        90       100       110       120
m040.pep    HGARHFLDRHAAAQGRTPHYCRGLRVTDETSLEQAQQFAGTVRSRFEAALCGSVSGFARA
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:.
a040        HGARHFLDRHAAAQGRTPHYCRGLRVTDETSLEQAQQFAGTVRSRFEAALCGSVSGFARA
                        70        80        90       100       110       120

                       130       140       150       160       170       180
m040.pep    PSVPLVSGNFLTARPIGVIDGTDMEYAGVIRKTDTAALRFQLDAGNIVWLPPLGHSYSGK
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a040        PSVPLVSGNFLTARPIGVIDGTDMEYAGVIRKTDTAALRFQLDAGNIVWLPPLGHSYSGK
                       130       140       150       160       170       180

                       190       200       210       220       230       240
m040.pep    TFYLDMLQTAASAAVSLQAEKLVYLTLSDGISRPDGTLAETLSAQEAQSLAEHAGGQTRR
            ||:|||||||||:|||||||||||||||||||||||||| ||||||||||||||||:|||
a040        TFHLDMLQTAASVAVSLQAEKLVYLTLSDGISRPDGTLAVTLSAQEAQSLAEHAGGETRR
```

```
                    190       200       210       220       230       240

                                                  250       260       270
m040.pep      LISSA------------------------ELFTRNGIGTSIAKEAFVSIRQAHXXDIPHI
              |||||                        |||||||||||:||||||||||||  |||||
a040          LISSAVAALEGGVHRVQILNGAADGSLLQELFTRNGIGTSIAKEAFVSIRQAHSGDIPHI
                    250       260       270       280       290       300

               280       290       300       310       320       330
m040.pep      AALIRPLEEQGILLHRXREYLENHISEFSILEHDGNLYGCAALKTFAEADCGEIACLAVS
              |||||||||||||||| |||||||||||||| |::|||||||||||||||||||||||||
a040          AALIRPLEEQGILLHRSREYLENHISEFSILEHDGNLYGCAALKTFAEADCGEIACLAVS
                    310       320       330       340       350       360

               340       350       360       370       380       390
m040.pep      PQXQDGGYGERXLAHIIDKARGIGISRLFALSTNTGEWFAERGFQTASEDELPETRRKDY
              || |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
a040          PQAQDGGYGERLLAHIIDKARGIGISRLFALSTNTGEWFAERGFQTASEDELPETRRKDY
                    370       380       390       400       410       420

               400       410
m040.pep      RSNGRNSHILVRRLHRX
              |||||||||||||||||
a040          RSNGRNSHILVRRLHRX
                    430
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 040 shows 88.3% identity over a 436 aa overlap with a predicted ORF (ORF 040.ng) from *N. gonorrhoeae:*

m040/g040

```
m040.pep      MSAPDLFVAHFREAVPYIRQMRGKTLVAGIDDRLLEGDTLNKLAADIGLLSQLGIRLVLI   60
              |:||| ||||||||:||||||| ||||||| ||||  ||||||||||||||||||||||
g040          MNAPDSFVAHFREAAPYIRQMRGTTLVAGIDGRLLEGGTLNKLAADIGLLSQLGIRLVLI   60

m040.pep      HGARHFLDRHAAAQGRTPHYCRGLRVTDETSLEQAQQFAGTVRSRFEAALCGSVSGFARA  120
              ||| ||||| |||||||||||||||||||||| ||||||||||||||||||||||||||
g040          HGAYHFLDRLAAAQGRTPHYCRGLRVTDETSLGQAQQFAGTVRSRFEAALCGSVSGFARA  120

m040.pep      PSVPLVSGNFLTARPIGVIDGTDMEYAGVIRKTDTAALRFQLDAGNIVWLPPLGHSYSGK  180
              ||||||||||||||||:|||||||||||||||||||||||||||||||||:||||||||:||
g040          PSVPLVSGNFLTARPMGVIDGTDMEYAGVIRKTDTAALRFQLDAGNIVWMPPLGHSYGGK  180

m040.pep      TFYLDMLQTAASAAVSLQAEKLVYLTLSDGISRPDGTLAETLSAQEAQSLAEHAGGQTRR  240
              || |||:|:|||:||||||||||||||||||||||||||||||||||||||||:::|||
g040          TFNLDMVQAAASVAVSLQAEKLVYLTLSDGISRPDGTLAETLSAQEAQSLAEHAASETRR  240

m040.pep      LISSA------------------------ELFTRNGIGTSIAKEAFVSIRQAHXXDIPHI  276
              |||||                        ||||||||||||||||||||||||  |||||
g040          LISSAVAALEGGVHRVQILNGAADGSLLQELFTRNGIGTSIAKEAFVSIRQAHSGDIPHI  300

m040.pep      AALIRPLEEQGILLHRXREYLENHISEFSILEHDGNLYGCAALKTFAEADCGEIACLAVS  336
              |||||||||||:|||| ||||||||||||||||||||:||||||||||||||||||||||
g040          AALIRPLEEQGVLLHRSREYLENHISEFSILEHDGDLYGCAALKTFAEADCGEIACLAVS  360

m040.pep      PQXQDGGYGERXLAHIIDKARGIGISRLFALSTNTGEWFAERGFQTASEDELPETRRKDY  396
              || |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
g040          PQAQDGGYGERLLAHIIDKARG_'ISRLFALSTNTGEWFAERGFQTASEDELPETRRKDY  420

m040.pep      RSNGRNSHILVRRLHRX  413
              |||||| ||||||||||
g040          RSNGRNPHILVRRLHRX  437
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 153>:

```
g041.seq
      1  ATGAGTTCGC CCAAACACAT CGGCTTGCAG GGCGGCAGCA ACGGCGGCCT
     51  GATTACCGCC GCCGCCTTCG TGCGCGAACC GCAAAGCATC GGTGCGCTGG
    101  TGTGCGAAGT ACCGCTGACC GATATGATCC GTTATCCGCT GCTGTCCGCC
    151  GGTTCAAGTT GGACGGACGA ATACGGCAAT CCGCAGAAAT ACGAAGCCTG
    201  CAAACGCCGG CTGGGCGAAT TGTCGCCGTA TCACAATCTT TCAGACGGCA
    251  TCGATTATCC GCCCGCACTC ATTACCACCA GCCTCAGCGA CGACCGCGTC
    301  CATCCCGCCC ACGCGCTCAA ATTCTACGCC AAACTGCGCG AAACCTCGCC
    351  GCAATCTTGG CTCTACTCGC CTGACGGCGG CGGCCATACC GGCAACGGCA
    401  CCCAACGCGA ATCCGCCGAC AAACTCGCCT GCGTGTTGCT GTTTTTGAAA
    451  GAATTTTTGG GATAA
```

This corresponds to the amino acid sequence <SEQ ID 154; ORF 041.ng>:

```
g041.pep
      1  MSSPKHIGLQ GGSNGGLITA AAFVREPQSI GALVCEVPLT DMIRYPLLSA
     51  GSSWTDEYGN PQKYEACKRR LGELSPYHNL SDGIDYPPAL ITTSLSDDRV
    101  HPAHALKFYA KLRETSPQSW LYSPDGGGHT GNGTQRESAD KLACVLLFLK
    151  EFLG*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 155>:

```
m041.seq
      1  ATCAGTTCGC CCGAACACAT CGGCTTGCAG GGCGGCAGCA ACGGCGGACT
     51  GATTACTGCC GCCGCCTTCG TGCGCGAACC GCAAAGCATC GGCGCGCTGG
    101  TGTGCGAAGT GCCGCTGACC GACATGATCC GTTATCCGCT GCTCTCCGCC
    151  GGTTCAAGCT GGACAGACGA ATACGGCAAT CCGCAAAAAT ACGAAGTCTG
    201  CAAACGCCGG TTGGGCGAAT TGTCGCCGTA TCACAATCTT TCAGACGGCA
    251  TCGATTATCC GCCCGCGCTC ATTACCACCA GCCTGTCCGA CGATCGCGTC
    301  CATCCCGCCC ACGCGCTCAA GTTCTACGCC AAACTGCGCG AAACCTCCGC
    351  GCAATCTTGG CTCTACTCGC CTGACGGCGG CGGCCATACC GGCAACGGCA
    401  CCCAACGCGA ATCCGCCGAC GAACTCGCCT GCGTCTTGCT GTTTTTGAAA
    451  GAGTTTTTGG GCTAA
```

This corresponds to the amino acid sequence <SEQ ID 156; ORF 041>:

```
m041.pep
      1  ISSPEHIGLQ GGSNGGLITA AAFVREPQSI GALVCEVPLT DMIRYPLLSA
     51  GSSWTDEYGN PQKYEVCKRR LGELSPYHNL SDGIDYPPAL ITTSLSDDRV
    101  HPAHALKFYA KLRETSAQSW LYSPDGGGHT GNGTQRESAD ELACVLLFLK
    151  EFLG*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 157>:

```
a041.seq
      1  ATCAGTTCGC CCGAACACAT CGGCTTGCAG GGCGGCAGCA ACGGCGGACT
     51  GATTACTGCC GCCGCCTTCG TGCGCGAACC GCAAAGCATA GGCGCGCTGG
    101  TGTGCGAAGT GCCGCTGACC GACATGATCC GTTATCCGCT GCTCTCCGCC
    151  GGTTCAAGCT GGACAGACGA ATACGGCAAT CCGCAAAAAT ACGAAGTCTG
    201  CAAACGCCGG TTGGGCGAAT TGTCGCCGTA TCACAATCTT TCAGACGGCA
    251  TCGATTATCC GCCCGCGCTC ATTACCACCA GCCTGTCCGA CGATCGCGTC
    301  CATCCCGCCC ACGCGCTCAA GTTCTACGCC AAACTGCGCG AAACCTCGCC
    351  GCAATCTTGG CTCTACTCGC CTGACGGCGG CGGCCATACC GGCAACGGCA
    401  CGCAGCGCGA AGCCGCCGAC GAACTCGCCT GCGTGTTGCT GTTTTTGAAA
    451  GAGTTTTTGG GCTAA
```

This corresponds to the amino acid sequence <SEQ ID 158; ORF 041.a>:

```
a041.pep
      1  ISSPEHIGLQ GGSNGGLITA AAFVREPQSI GALVCEVPLT DMIRYPLLSA
     51  GSSWTDEYGN PQKYEVCKRR LGELSPYHNL SDGIDYPPAL ITTSLSDDRV
    101  HPAHALKFYA KLRETSPQSW LYSPDGGGHT GNGTQREAAD ELACVLLFLK
```

```
151  EFLG*
```

m041/a041 98.7% identity over a 154 aa overlap

```
                    10        20        30        40        50        60
m041.pep    ISSPEHIGLQGGSNGGLITAAAFVREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGN
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a041        ISSPEHIGLQGGSNGGLITAAAFVREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGN
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m041.pep    PQKYEVCKRRLGELSPYHNLSDGIDYPPALITTSLSDDRVHPAHALKFYAKLRETSAQSW
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
a041        PQKYEVCKRRLGELSPYHNLSDGIDYPPALITTSLSDDRVHPAHALKFYAKLRETSPQSW
                    70        80        90       100       110       120

                   130       140       150
m041.pep    LYSPDGGGHTGNGTQRESADELACVLLFLKEFLGX
            ||||||||||||||||||:||||||||||||||||
a041        LYSPDGGGHTGNGTQREAADELACVLLFLKEFLGX
                   130       140       150
```

Computer analysis of this amino acid sequence gave the following results:
<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>
ORF 041 shows 96.8% identity over a 154 aa overlap with a predicted ORF (ORF 041.ng) from *N. gonorrhoeae:*
m041/g041

```
                    10        20        30        40        50        60
m041.pep    ISSPEHIGLQGGSNGGLITAAAFVREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGN
            :|||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
g041        MSSPKHIGLQGGSNGGLITAAAFVREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGN
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m041.pep    PQKYEVCKRRLGELSPYHNLSDGIDYPPALITTSLSDDRVHPAHALKFYAKLRETSAQSW
            |||||:||||||||||||||||||||||||||||||||||||||||||||||||| ||||
g041        PQKYEACKRRLGELSPYHNLSDGIDYPPALITTSLSDDRVHPAHALKFYAKLRETSPQSW
                    70        80        90       100       110       120

                   130       140       150
m041.pep    LYSPDGGGHTGNGTQRESADELACVLLFLKEFLGX
            ||||||||||||||||||||:||||||||||||||
g041        LYSPDGGGHTGNGTQRESADKLACVLLFLKEFLGX
                   130       140       150
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 159>:

```
g041-1.seq
      1  ATGAAATCCT ACCCCGACCC CTACCGCCAT TTTGAAAACC TCGATTCCGC
     51  CGAAACGCAA AACTTCGCTG CTGAAGCGAA TGCCGAAACG CGCGCGCGTT
    101  TTTTAAACAA CGACAAGGCG CGCGCACTTT CAGACGGCAT TTTGAATCAA
    151  ATGCAGGACA CGCGGCAGAT TCCGTTTTGT CAGGAACACC GCGCGCGGAT
    201  GTACCATTTC CATCAGAATG CGGAATATCC GAAGGGCGTG TACCGCATGT
    251  GTACGGCGGC GACCTACCGT TCCGGCTATC CCGAGTGGAA AATCCTGTTT
    301  TCGGTGGCGG ATTTCGATGA GTTGCTCGGC GACGATGTGT ATTTGGGCGG
    351  CGTGTCGCAC TTGGTGGAGC AGCCCAACCG CGCGCTGCTG ACTTTGAACA
    401  AATCGGGCGG CGATACGGCG TATACGCTGG AAGTGGATTT GGAAGCAGGG
    451  GAATTGGTAG AGGGCGGTTT TCACTTTCCG GCAGGCAAAA ACCATGTGTC
    501  GTGGCGCGAT GAAAACAGCG TGTGGGTGTG TCCGGCTTGG GACGAACGCC
    551  AGTTGACCGA ATCGGGCTAT CCGCGCGAAG TGTGGCTGGT GGAACGCGGC
```

```
 601  AAGAGTTTCG AGGAAAGCCT GCCGGCGTAC CAAATCGATA AAGGCGCGAT
 651  GATGGTAAAC GCGTGGCGTT ACCTCGATCC GCAGGGTTCG CCGATTGATT
 701  TGATTGAAGC GTCGGACGGT TTTTACACCA AGACGTATTT GCAGGTGTCG
 751  TCCGAAGGCG GGGCGAAACC GTTGAACCTG CCTAATGATT GCGATGTGGT
 801  CGGCTATCTG GCGGGACATC TTTTGCTGAC GCTGCGCAAG GACTGGCACC
 851  GCGCGAACCA AAGCTATCCG AGTGGCGCGT TGGTGGCGGT GAAACTGAAT
 901  CGGGGCGAAC TCGGGGCGGC GCAGCTTTTG TTTGCGCCCG ATGAAACGCA
 951  GGCATTGGAA AGCGTGGAAA CGACCAAGCG TTTTGTGGTG GCAAGCCTGC
1001  TGGAGAATGT ACAAGGCCGT CTGAAAGCGT GGCGGTTTGC CGACAGCAAA
1051  TGGCAGGAAG CCGAGTTGCC GCACCTGCCC TCGGGCGCGT TGGAAATGAC
1101  CGACCAACCG TGGGGCGGCG ACGTGGTTTA TCTTGCCGCC AGCGATTTCA
1151  CCACGCCGCT GACGCTGTTT GCGCTGGATT TGAACGTGAT GGAACTGACC
1201  GTCATGCGCC TCCAGCCGCA GCAGTTTGTT TCAGACGGCA TCGAAGTGCG
1251  GCAGTTTTGG GCGGTGTCGT CCGACGGCGA ACGCATTCCT TATTTCCACG
1301  TCGGCAAAAA CGCCGCGCCC GACACGCCGA CCTTAGTCTA TGCTTACGGA
1351  GGTTTCGGCA TTCCTGAATT GCCGCATTAT CTGGGCAGCG TCGGCAAATA
1401  TTGGCTGGAA GAGGGCAATG CCTTTGTATT GGCAAACATC CGCGGCGGCG
1451  GAGAATTCGG CCCGCGCTGG CATCAGGCGG CGCAGGGAAT CAGCAAACAC
1501  AAAAGCGTTG ATGATTTGTT GGCAGTCGTG CGTGATTTGT CCGAACGCGG
1551  CATGAGTTCG CCCAAACACA TCGGCTTGCA GGGCGGCAGC AACGGCGGCC
1601  TGATTACCGC CGCCGCCTTC GTGCGCGAAC CGCAAAGCAT CGGTGCGCTG
1651  GTGTGCGAAG TACCGCTGAC CGATATGATC CGTTATCCGC TGCTGTCCGC
1701  CGGTTCAAGT TGGACGGACG AATACGGCAA TCCGCAGAAA TACGAAGCCT
1751  GCAAACGCCG GCTGGGCGAA TTGTCGCCGT ATCACAATCT TTCAGACGGC
1801  ATCGATTATC CGCCCGCACT CATTACCACC AGCCTCAGCG ACGACCGCGT
1851  CCATCCCGCC CACGCGCTCA AATTCTACGC CAAACTGCGC GAAACCTCGC
1901  CGCAATCTTG GCTCTACTCG CCTGACGGCG GCGGCCATAC CGGCAACGGC
1951  ACCCAACGCG AATCCGCCGA CAAACTCGCC TGCGTGTTGC TGTTTTTGAA
2001  AGAATTTTTG GGATAA
```

This corresponds to the amino acid sequence <SEQ ID 160; ORF 041-1.ng>:

```
g041-1.pep
   1  MKSYPDPYRH FENLDSAETQ NFAAEANAET RARFLNNDKA RALSDGILNQ
  51  MQDTRQIPFC QEHRARMYHF HQNAEYPKGV YRMCTAATYR SGYPEWKILF
 101  SVADFDELLG DDVYLGGVSH LVEQPNRALL TLNKSGGDTA YTLEVDLEAG
 151  ELVEGGFHFP AGKNHVSWRD ENSVWVCPAW DERQLTESGY PREVWLVERG
 201  KSFEESLPAY QIDKGAMMVN AWRYLDPQGS PIDLIEASDG FYTKTYLQVS
 251  SEGGAKPLNL PNDCDVVGYL AGHLLLTLRK DWHRANQSYP SGALVAVKLN
 301  RGELGAAQLL FAPDETQALE SVETTKRFVV ASLLENVQGR LKAWRFADSK
 351  WQEAELPHLP SGALEMTDQP WGGDVVYLAA SDFTTPLTLF ALDLNVMELT
 401  VMRLQPQQFV SDGIEVRQFW AVSSDGERIP YFHVGKNAAP DTPTLVYAYG
 451  GFGIPELPHY LGSVGKYWLE EGNAFVLANI RGGGEFGPRW HQAAQGISKH
 501  KSVDDLLAVV RDLSERGMSS PKHIGLQGGS NGGLITAAAF VREPQSIGAL
 551  VCEVPLTDMI RYPLLSAGSS WTDEYGNPQK YEACKRRLGE LSPYHNLSDG
 601  IDYPPALITT SLSDDRVHPA HALKFYAKLR ETSPQSWLYS PDGGGHTGNG
 651  TQRESADKLA CVLLFLKEFL G*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 161>:

```
m041-1.seq
   1  ATGAAATCCT ACCCCGACCC CTACCGCCAT TTTGAAAACC TCGATTCCGC
  51  CGAAACGCAA AACTTCGCTG CTGAAGCGAA TGCCGAAACG CGCGCGCGTT
 101  TTTTAGAAAA CGACAAGGCG CGCGCGCTTT CAGACGGCAT TTTGGCGCAG
 151  TTGCAGGACA CGCGGCAGAT TCCGTTTTGT CAGGAACACC GCGCGCGGAT
 201  GTACCATTTC CATCAGGACG CGGAGTATCC GAAGGGCGTG TACCGCGTGT
 251  GTACCGCGGC GACGTATCGT TCCGGCTATC CCGAGTGGAA AATCCTGTTT
 301  TCGGTGGCGG ATTTCGACGA ATTGCTTGGC GACGATGTGT ATTTGGGCGG
 351  CGTGTCGCAC TTGGTGGAAC AGCCCAACCG CGCGTTGTTA ACACTGAGCA
 401  AATTGGGCAG CGATACGGCG TACACGCTGG AAGTGGATTT GGAAGCAGGG
 451  GAGTTGGTCG AAGGCGGTTT TCACTTTCCG GCAGGCAAAA ACCATGTGTC
 501  GTGGCGCGAT GAAAACAGCG TGTGGGTGTG TCCGGCTTGG AACGAACGCC
 551  AGTTGACCCA ATCGGGCTAT CCGCGCGAAG TATGGCTGGT GGAACGCGGC
 601  AAGAGTTTCG AGGAAAGCCT GCCTGTGTAT CAAATCGGCG AAGACGGCAT
 651  GATGGTGAAC GCGTGGCGTT ATCTCGATCC GCAGGGTTCG CCGATTGATT
 701  TGATTGAAGC GTCGGACGGT TTTTACACCA AAACCTATTT GCGGGTCTCA
```

```
 751  GCCGAAGGCG AGGCGAAACC GTTAAACCTG CCCAACGATT GCGACGTGGT
 801  CGGCTATCTG GCGGGGCATC TTTTGCTGAC GCTGCGCAAG GACTGGAACC
 851  GCGCGAACCA AAGCTATCCG AGCGGCGCGC TGGTGGCGGT GAAGCTGAAT
 901  CGGGGCGAAC TCGGGGCGGC GCAGCTTTTG TTTGCGCCCG ATGAAACGCA
 951  GGCATTGGAA AGCGTGGAAA CGACCAAGCG TTTTGTGGTG GCGAGCCTGT
1001  TGGAGAACGT ACAAGGCCGT CTGAAAGCAT GGCGGTTTGC CGACGGCAAA
1051  TGGCAGGAAG TCGAATTGCC GCGCCTGCCT TCGGGCGCGT TGGAAATGAC
1101  CGACCAACCT TGGGGCGGCG ACGTGGTTTA CCTTGCCGCC AGCGATTTCA
1151  CCACGCCGCT GACGCTGTTT GCGCTGGATT TGAACGTGAT GGAACTGACC
1201  GTCATGCGCC GCCAGCCGCA GCAGTTTGAT TCAGACGGCA TTAACGTGCA
1251  GCAGTTTTGG ACGACTTCGG CTGACGGCGA GCGCATTCCT TATTTCCACG
1301  TCGGCAAAAA CGCCGCGCCC GACATGCCGA CCGCGGTCTA TGCCTACGGC
1351  GGTTTCGGCA TTCCCGAATT GCCGCATTAT CTGGGCAGCA TTGGCAAATA
1401  TTGGCTGGAA GAGGGCAATG CCTTTGTATT GGCGAACATC CGCGGCGGCG
1451  GCGAGTTCGG CCCGCGCTGG CATCAGGCGG CGCAGGGAAT CAGCAAACAT
1501  AAAAGCGTTG ATGATTTATT GGCAGTCGTG CGCGATTTGT CCGAACGCGG
1551  TATCAGTTCG CCCGAACACA TCGGCTTGCA GGGCGGCAGC AACGGCGGAC
1601  TGATTACTGC CGCCGCCTTC GTGCGCGAAC CGCAAAGCAT CGGCGCGCTG
1651  GTGTGCGAAG TGCCGCTGAC CGACATGATC CGTTATCCGC TGCTCTCCGC
1701  CGGTTCAAGC TGGACAGACG AATACGGCAA TCCGCAAAAA TACGAAGTCT
1751  GCAAACGCCG GTTGGGCGAA TTGTCGCCGT ATCACAATCT TTCAGACGGC
1801  ATCGATTATC CGCCCGCGCT CATTACCACC AGCCTGTCCG ACGATCGCGT
1851  CCATCCCGCC CACGCGCTCA AGTTCTACGC CAAACTGCGC GAAACCTCCG
1901  CGCAATCTTG GCTCTACTCG CCTGACGGCG GCGGCCATAC CGGCAACGGC
1951  ACCCAACGCG AATCCGCCGA CGAACTCGCC TGCGTCTTGC TGTTTTTGAA
2001  AGAGTTTTTG GGCTAA
```

This corresponds to the amino acid sequence <SEQ ID 162; ORF 041-1>:

m041-1.pep

```
  1  MKSYPDPYRH FENLDSAETQ NFAAEANAET RARFLENDKA RALSDGILAQ
 51  LQDTRQIPFC QEHRARMYHF HQDAEYPKGV YRVCTAATYR SGYPEWKILF
101  SVADFDELLG DDVYLGGVSH LVEQPNRALL TLSKLGSDTA YTLEVDLEAG
151  ELVEGGFHFP AGKNHVSWRD ENSVWVCPAW NERQLTQSGY PREVWLVERG
201  KSFEESLPVY QIGEDGMMVN AWRYLDPQGS PIDLIEASDG FYTKTYLRVS
251  AEGEAKPLNL PNDCDVVGYL AGHLLLTLRK DWNRANQSYP SGALVAVKLN
301  RGELGAAQLL FAPDETQALE SVETTKRFVV ASLLENVQGR LKAWRFADGK
351  WQEVELPRLP SGALEMTDQP WGGDVVYLAA SDFTTPLTLF ALDLNVMELT
401  VMRRQPQQFD SDGINVQQFW TTSADGERIP YFHVGKNAAP DMPTLVYAYG
451  GFGIPELPHY LGSIGKYWLE EGNAFVLANI RGGGEFGPRW HQAAQGISKH
501  KSVDDLLAVV RDLSERGISS PEHIGLQGGS NGGLITAAAF VREPQSIGAL
551  VCEVPLTDMI RYPLLSAGSS WTDEYGNPQK YEVCKRRLGE LSPYHNLSDG
601  IDYPPALITT SLSDDRVHPA HALKFYAKLR ETSAQSWLYS PDGGGHTGNG
651  TQRESADELA CVLLFLKEFL G*
```

m041-1/g041-1 94.6% identity in 671 aa overlap

```
                    10        20        30        40        50        60
m041-1.pep  MKSYPDPYRHFENLDSAETQNFAAEANAETRARFLENDKARALSDGILAQLQDTRQIPFC
            |||||||||||||||||||||||||||||||||||:||||||||||||| |:|||||||||
g041-1      MKSYPDPYRHFENLDSAETQNFAAEANAETRARFLNNDKARALSDGILNQMQDTRQIPFC
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m041-1.pep  QEHRARMYHFHQDAEYPKGVYRVCTAATYRSGYPEWKILFSVADFDELLGDDVYLGGVSH
            ||||||||||||:|||||||||:|||||||||||||||||||||||||||||||||||||
g041-1      QEHRARMYHFHQNAEYPKGVYRMCTAATYRSGYPEWKILFSVADFDELLGDDVYLGGVSH
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m041-1.pep  LVEQPNRALLTLSKLGSDTAYTLEVDLEAGELVEGGFHFPAGKNHVSWRDENSVWVCPAW
            ||||||||||||:| |:|||||||||||||||||||||||||||||||||||||||||||
g041-1      LVEQPNRALLTLNKSGGDTAYTLEVDLEAGELVEGGFHFPAGKNHVSWRDENSVWVCPAW
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m041-1.pep  NERQLTQSGYPREVWLVERGKSFEESLPVYQIGEDGMMVNAWRYLDPQGSPIDLIEASDG
            :|||||:||||||||||||||||||||:||| : :|||||||||||||||||||||||||
```

```
g041-1        DERQLTESGYPREVNLVERGKSFEESLPAYQIDKGAMMVNAWRYLDPQGSPIDLIEASDG
                      190       200       210       220       230       240

                      250       260       270       280       290       300
m041-1.pep    FYTKTYLRVSAEGEAKPLNLPNDCDVVGYLAGHLLLTLRKDWNRANQSYPSGALVAVKLN
              |||||||:||:|| ||||||||||||||||||||||||||||:|||||||||||||||||
g041-1        FYTKTYLQVSSEGGAKPLNLPNDCDVVGYLAGHLLLTLRKDWHRANQSYPSGALVAVKLN
                      250       260       270       280       290       300

                      310       320       330       340       350       360
m041-1.pep    RGELGAAQLLFAPDETQALESVETTKRFVVASLLENVQGRLKAWRFADGKWQEVELPRLP
              ||||||||||||||||||||||||||||||||||||||||||||||||:||||:|||:||
g041-1        RGELGAAQLLFAPDETQALESVETTKRFVVASLLENVQGRLKAWRFADSKWQEAELPHLP
                      310       320       330       340       350       360

                      370       380       390       400       410       420
m041-1.pep    SGALEMTDQPWGGDVVYLAASDFTTPLTLFALDLNVMELTVMRRQPQQFDSDGINVQQFW
              ||||||||||||||||||||||||||||||||||||||||||| ||||| ||||:|:|||
g041-1        SGALEMTDQPWGGDVVYLAASDFTTPLTLFALDLNVMELTVMRLQPQQFVSDGIEVRQFW
                      370       380       390       400       410       420

                      430       440       450       460       470       480
m041-1.pep    TTSADGERIPYFHVGKNAAPDMPTLVYAYGGFGIPELPHYLGSIGKYWLEEGNAFVLANI
              ::|:||||||||||||||||| |||||||||||||||||||||:||||||||||||||||
g041-1        AVSSDGERIPYFHVGKNAAPDTPTLVYAYGGFGIPELPHYLGSVGKYWLEEGNAFVLANI
                      430       440       450       460       470       480

                      490       500       510       520       530       540
m041-1.pep    RGGGEFGPRWHQAAQGISKHKSVDDLLAVVRDLSERGISSPEHIGLQGGSNGGLITAAAF
              |||||||||||||||||||||||||||||||||||||:|||:||||||||||||||||||
g041-1        RGGGEFGPRWHQAAQGISKHKSVDDLLAVVRDLSERGMSSPKHIGLQGGSNGGLITAAAF
                      490       500       510       520       530       540

                      550       560       570       580       590       600
m041-1.pep    VREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGNPQKYEVCKRRLGELSPYHNLSDG
              ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
g041-1        VREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGNPQKYEACKRRLGELSPYHNLSDG
                      550       560       570       580       590       600

                      610       620       630       640       650       660
m041-1.pep    IDYPPALITTSLSDDRVHPAHALKFYAKLRETSAQSWLYSPDGGGHTGNGTQRESADELA
              |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||:||
g041-1        IDYPPALITTSLSDDRVHPAHALKFYAKLRETSPQSWLYSPDGGGHTGNGTQRESADKLA
                      610       620       630       640       650       660

                      670
m041-1.pep    CVLLFLKEFLGX
              ||||||||||||
g041-1        CVLLFLKEFLGX
                      670
```

m041-1/P55577

```
sp|P55577|Y4NA_RHISN PROBABLE PEPTIDASE Y4NA >gi|2182536 (AE000086) Y4nA [Rhizobium sp.
NGR234] Length = 726
 Score =  370 bits (940), Expect = e-101
 Identities = 217/682 (31%), Positives = 331/682 (47%), Gaps = 22/682 (3%)

Query: 2    KSYPDPYRHFENLDSAETQNFAAEANAETRARFLENDKARALSDGILAQLQDTRQIPFCQ 61
            K   DP  +   +D  +   +    N  T  +  ++ +          L  LQ T +I
Sbjct: 42   KDASDPRAYLNEIDGDKAMTWVEAHNLSTVDKLSKDPRYSEYQADALTILQATDRIASPS 101

Query: 62   EHRARMY-HFHQDAEYPKGVYRVCTAATYRSGYPEWKILFSVADFDELLGDDVYLGGVSH 120
              R  M  +F QD  + +G++R  T  +YRSG P+W+ +  V     +  G     G
Sbjct: 102  FARDGMIDNFWQDGTHVQGLWRRTTWESYRSGNPQWRTILDVDALSKAEGKTWVFEGGDC 161

Query: 121  LVEQPNRALLTLSKLGSDTAYTLEVDLEAGELVEGGFHFPAGKNHVSWRDENSVWVCPAW 180
            L    N  L+ LS  G D      E D+  GE V+ GF  P GK  V+W DEN+++V   W
```

```
Sbjct: 162 LPPTSNLCLIRLSDGGKDADVVREFDIAKGEFVKEGFVLPEGKQSVTWVDENTIYVTREW 221

Query: 181 NERQLTQSGYPREVWLVERGKSFEESLPVYQ------IGEDGMM--VNAWRYLDPQGSPI 232
              ++T SGY     +V+RG+S ++++ +++          E G++  ++     +D     +
Sbjct: 222 TPGEVTSSGYAYVTKVVKRGQSLDQAVEIFRGQKKDVSAERGVLRDIDGKYVMDTSYRGL 281

Query: 233 DLIEASDGFYTKTYLRVSAEGEAKPLNLPNDCDVVGYLAGHLLLTLRKDWNRANQS-YPS 291
            D       FY    +       +     L LP         GY  G   +  L+ DW   A   + + +
Sbjct: 282 DFFNTELAFYPNGH----PDTRKVVLPLPTTAVFSGYYKGQAIYWLKSDWTSAKGTVFHN 337

Query: 292 GALVAVKLNRGELGAAQL----LFAPDETQALESVETTKRFVVASLLENVQGRLKAWRFA 347
            GA++A  L        A++     LF P+E Q++         TK  +V S+L NV      ++++ F
Sbjct: 338 GAIIAFDLKAALADPARVEPLVLFMPNEHQSVAGTTQTKNRLVLSILLNVTSEVRSFDFG 397

Query: 348 DGKWQEVELPRLPSGALEMTDQPWGGDVVYLAASDFTTPLTLFALDLNVMELTVMRRQPQ 407
             G W   +L     +  L +T        D +++ +   F  P TLF  D      ++   +   P
Sbjct: 398 KGGWSSFKLALPENSTLSLTSSDDESDQLFVFSEGFLEPSTLFCADAATGQVEKITSTPA 457

Query: 408 QFDSDGINVQQFWTTSADGERIPYFHVGKNAAP---DMPTLVYAYGGFGIPELPHYLGSI 464
            +FD+ G+   QQFW TS DG ++PYF V +             PT++YAYGGF IP   P Y   +
Sbjct: 458 RFDAGGLQAQQFWATSKDGTKVPYFLVARKDVKLDGTNPTILYAYGGFQIPMQPSYSAVL 517

Query: 465 GKYWLEEGNAFVLANIRGGGEFGPRWHQAAQGISKHKSVDDLLAVVRDLSERGISSPEHI 524
            GK WLE+G A+ LANIRGGGEFGP+WH A     ++ +  DD  AV +DL  + ++S  H+
Sbjct: 518 GKLWLEKGGAYALANIRGGGEFGPKWHDAGLKTNRQRVYDDFQAVAQDLIAKKVTSTPHL 577

Query: 525 GLQGGSNGGLITAAAFVREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGNPQKYEVC 584
            G+ GGSNGGL+       ++ P     A+V +VPL DM+ +   +SAG+SW  EYG+P     V
Sbjct: 578 GIMGGSNGGLLMGVQMIQRPDLWNAVVIQVPLLDMVNFTRMSAGASWQAEYGSPDD-PVE 636

Query: 585 KRRLGELSPYHNLSDGIDYPPALITTSLSDDRVHPAHALKFYAKLRETSAQSWLYSPDGG 644
               L  +SPYHN+  G+ YP       TS  DDRV P HA K  A    +      + Y    G
Sbjct: 637 GAFLRSISPYHNVKAGVAYPEPFFETSTKDDRVGPVHARKMAALFEDMGLPFYYYENIEG 696

Query: 645 GHTGNGTQRESADELACVLLFL 666
            GH      +E A   A   +++
Sbjct: 697 GHAAAANLQEHAARLALEYIYM 718
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 163>:

```
a041-1.seq
      1  ATGAAATCCT ACCCCGACCC CTACCGCCAT TTTGAAAACC TCGATTCCGC
     51  CGAAACGCAA AACTTCGCTG CTGAAGCGAA TGCCGAAACG CGCGCGCGTT
    101  TTTTAAACAA CGACAAGGCA CGCGCATTGT CTGACGGCAT TTTGGCGCAG
    151  TTGCAGGACA CGCGGCAAAT TCCGTTTTGT CAGGAACACC GCGCGCGGAT
    201  GTACCATTTC CATCAAGATG CGGAATATCC GAAAGGCGTG TACCGCGTGT
    251  GTACCGCGGC GACTTACCGT TCGGGCTATC CTGAGTGGAA AATCCTGTTT
    301  TCGGTGGCGG ATTTCGACGA ATTGCTCGGT GACGATGTAT ATCTAGGCGG
    351  CGTGTCGCAC CTGGTGGAAC AGCCCAACCG CGCGTTGTTA ACACTGAGCA
    401  AATCGGGCGG CGATACCGCG TACACGCTGG AAGTGGATTT GGAAGCAGGG
    451  GAGTTGGTAG AAGGCGGTTT TCACTTTCCG GCAGGCAAAA ACCATGTGTC
    501  GTGGCGCGAT GAAAACAGCG TGTGGGTGTG TCCGGCTTGG GACGAACGCC
    551  AGTTGACCGA ATCGGGCTAT CCGCGCGAGG TGTGGCTGGT GGAACGCGGC
    601  AAGAGTTTCG AGGAAAGCCT GCCGGTGTAC CAAATTGCTG AAGACGGCAT
    651  GATGGTGAAC GCGTGGCGTT ACCTCGATCC GCAGGGTTCG CCGATTGATT
    701  TGATTGAAGC GTCTGACGGT TTTTACACCA AAACCTATTT GCAGGTCTCA
    751  GCCGAAGGCG AAGCGAAACC GTTAAACCTG CCCAACGATT GCGACGTAGT
    801  CGGCTATCTG GCCGGACATC TTTTGCTGAC CTTGCGTAAA GACTGGCACC
    851  GCGCGAACCA AAGCTATCCG AGTGGCGCAT TGGTAGCAGT AAAATTAAAC
    901  CGCGGCGAAT TGGGCGCGGC GCAGCTTTTG TTTGCGCCCA ATGAAACGCA
    951  GGCATTGGAA AGCGTGGAAA CGACCAAGCG TTTTGTCGTG GCGAGCCTGC
   1001  TGGAAAACGT ACAGGGTCGT CTGAAAGCGT GGCGTTTTAC TGATGGCAAA
   1051  TGGCAGGAAA CCGAGTTGCC GCGCCTGCCT TCGGGCGCGT TGGAAATGAC
   1101  CGACCAACCG TGGGGGGGCG ACGTAGTTTA CCTTGCCGCC AGCGATTTCA
   1151  CCACGCCGCT GACGCTGTTT GCATTGGATT TGAACGTGAT GGAACTGACC
   1201  GTCATGCGCC GCCAGCCGCA GCAGTTTGAT TCAGACGGCA TTAACGTGCA
   1251  GCAGTTTTGG ACGACTTCGG CTGACGGCGA GCGCATTCCT TATTTCCACG
   1301  TCGGCAAAAA CGCCGCGCCC GACATGCCGA CGCTGGTCTA TGCCTACGGC
   1351  GGTTTCGGCA TTCCCGAATT GCCGCATTAT CTGGGCAGCA TTGGCAAATA
   1401  TTGGCTGGAA GAGGGCAATG CCTTTGTATT GGCGAACATC CGCGGCGGCG
```

```
1451  GCGAGTTCGG CCCGCGCTGG CATCAGGCGG CGCAGGGAAT CAGCAAACAT
1501  AAAAGCGTTG ATGATTTATT GGCAGTCGTG AGCGATTTGT CCGAACGCGG
1551  TATCAGTTCG CCCGAACACA TCGGCTTGCA GGGCGGCAGC AACGGCGGAC
1601  TGATTACTGC CGCCGCCTTC GTGCGCGAAC CGCAAAGCAT AGGCGCGCTG
1651  GTGTGCGAAG TGCCGCTGAC CGACATGATC CGTTATCCGC TGCTCTCCGC
1701  CGGTTCAAGC TGGACAGACG AATACGGCAA TCCGCAAAAA TACGAAGTCT
1751  GCAAACGCCG GTTGGGCGAA TTGTCGCCGT ATCACAATCT TTCAGACGGC
1801  ATCGATTATC CGCCCGCGCT CATTACCACC AGCCTGTCCG ACGATCGCGT
1851  CCATCCCGCC CACGCGCTCA AGTTCTACGC CAAACTGCGC GAAACCTCGC
1901  CGCAATCTTG GCTCTACTCG CCTGACGGCG GCGGCCATAC CGGCAACGGC
1951  ACGCAGCGCG AAGCCGCCGA CGAACTCGCC TGCGTGTTGC TGTTTTTGAA
2001  AGAGTTTTTG GGCTAA
```

This corresponds to the amino acid sequence <SEQ ID 164; ORF 041-1.a>:

a041-1.pep

```
  1  MKSYPDPYRH FENLDSAETQ NFAAEANAET RARFLNNDKA RALSDGILAQ
 51  LQDTRQIPFC QEHRARMYHF HQDAEYPKGV YRVCTAATYR SGYPEWKILF
101  SVADFDELLG DDVYLGGVSH LVEQPNRALL TLSKSGGDTA YTLEVDLEAG
151  ELVEGGFHFP AGKNHVSWRD ENSVWVCPAW DERQLTESGY PREVWLVERG
201  KSFEESLPVY QIAEDGMMVN AWRYLDPQGS PIDLIEASDG FYTKTYLQVS
251  AEGEAKPLNL PNDCDVVGYL AGHLLLTLRK DWHRANQSYP SGALVAVKLN
301  RGELGAAQLL FAPNETQALE SVETTKRFVV ASLLENVQGR LKAWRFTDGK
351  WQETELPRLP SGALEMTDQP WGGDVVYLAA SDFTTPLTLF ALDLNVMELT
401  VMRRQPQQFD SDGINVQQFW TTSADGERIP YFHVGKNAAP DMPTLVYAYG
451  GFGIPELPHY LGSIGKYWLE EGNAFVLANI RGGGEFGPRW HQAAQGISKH
501  KSVDDLLAVV SDLSERGISS PEHIGLQGGS NGGLITAAAF VREPQSIGAL
551  VCEVPLTDMI RYPLLSAGSS WTDEYGNPQK YEVCKRRLGE LSPYHNLSDG
601  IDYPPALITT SLSDDRVHPA HALKFYAKLR ETSPQSWLYS PDGGGHTGNG
651  TQREAADELA CVLLFLKEFL G*
```

a041-1/m041-1 97.9% identity in 671 aa overlap

```
                    10        20        30        40        50        60
a041-1.pep  MKSYPDPYRHFENLDSAETQNFAAEANAETRARFLNNDKARALSDGILAQLQDTRQIPFC
            |||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
m041-1      MKSYPDPYRHFENLDSAETQNFAAEANAETRARFLENDKARALSDGILAQLQDTRQIPFC
                    10        20        30        40        50        60

                    70        80        90       100       110       120
a041-1.pep  QEHRARMYHFHQDAEYPKGVYRVCTAATYRSGYPEWKILFSVADFDELLGDDVYLGGVSH
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m041-1      QEHRARMYHFHQDAEYPKGVYRVCTAATYRSGYPEWKILFSVADFDELLGDDVYLGGVSH
                    70        80        90       100       110       120

                   130       140       150       160       170       180
a041-1.pep  LVEQPNRALLTLSKSGGDTAYTLEVDLEAGELVEGGFHFPAGKNHVSWRDENSVWVCPAW
            |||||||||||||| |:|||||||||||||||||||||||||||||||||||||||||||
m041-1      LVEQPNRALLTLSKLGSDTAYTLEVDLEAGELVEGGFHFPAGKNHVSWRDENSVWVCPAW
                   130       140       150       160       170       180

                   190       200       210       220       230       240
a041-1.pep  DERQLTESGYPREVWLVERGKSFEESLPVYQIAEDGMMVNAWRYLDPQGSPIDLIEASDG
            :|||||:|||||||||||||||||||||||||:|||||||||||||||||||||||||||
m041-1      NERQLTQSGYPREVWLVERGKSFEESLPVYQIGEDGMMVNAWRYLDPQGSPIDLIEASDG
                   190       200       210       220       230       240

                   250       260       270       280       290       300
a041-1.pep  FYTKTYLQVSAEGEAKPLNLPNDCDVVGYLAGHLLLTLRKDWHRANQSYPSGALVAVKLN
            |||||||:||||||||||||||||||||||||||||||||||:|||||||||||||||||
m041-1      FYTKTYLRVSAEGEAKPLNLPNDCDVVGYLAGHLLLTLRKDWNRANQSYPSGALVAVKLN
                   250       260       270       280       290       300

                   310       320       330       340       350       360
a041-1.pep  RGELGAAQLLFAPNETQALESVETTKRFVVASLLENVQGRLKAWRFTDGKWQETELPRLP
            |||||||||||||:||||||||||||||||||||||||||||||||||||:||||:||||
m041-1      RGELGAAQLLFAPDETQALESVETTKRFVVASLLENVQGRLKAWRFADGKWQEVELPRLP
                   310       320       330       340       350       360
```

```
                     370       380       390       400       410       420
a041-1.pep    SGALEMTDQPWGGDVVYLAASDFTTPLTLFALDLNVMELTVMRRQPQQFDSDGINVQQFW
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m041-1        SGALEMTDQPWGGDVVYLAASDFTTPLTLFALDLNVMELTVMRRQPQQFDSDGINVQQFW
                     370       380       390       400       410       420

                     430       440       450       460       470       480
a041-1.pep    TTSADGERIPYFHVGKNAAPDMPTLVYAYGGFGIPELPHYLGSIGKYWLEEGNAFVLANI
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m041-1        TTSADGERIPYFHVGKNAAPDMPTLVYAYGGFGIPELPHYLGSIGKYWLEEGNAFVLANI
                     430       440       450       460       470       480

                     490       500       510       520       530       540
a041-1.pep    RGGGEFGPRWHQAAQGISKHKSVDDLLAVVSDLSERGISSPEHIGLQGGSNGGLITAAAF
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
m041-1        RGGGEFGPRWHQAAQGISKHKSVDDLLAVVRDLSERGISSPEHIGLQGGSNGGLITAAAF
                     490       500       510       520       530       540

                     550       560       570       580       590       600
a041-1.pep    VREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGNPQKYEVCKRRLGELSPYHNLSDG
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m041-1        VREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGNPQKYEVCKRRLGELSPYHNLSDG
                     550       560       570       580       590       600

                     610       620       630       640       650       660
a041-1.pep    IDYPPALITTSLSDDRVHPAHALKFYAKLRETSPQSWLYSPDGGGHTGNGTQREAADELA
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||:|||||
m041-1        IDYPPALITTSLSDDRVHPAHALKFYAKLRETSAQSWLYSPDGGGHTGNGTQRESADELA
                     610       620       630       640       650       660

                     670
a041-1.pep    CVLLFLKEFLGX
              ||||||||||||
m041-1        CVLLFLKEFLGX
                     670
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 165>:

g042.seq

```
  1  ATGACGATGA TTTGCTTGCG CTTCCAagcG TTCGTGCCGC ATACCAGCGC
 51  GTTATCCAAC ACTTCCACGG CAGCCGGCCC TTCCTGCCCG ATGGCGGCGG
101  TGCGGTCGAT GATGAAAATC CAGCCGGGGT TTTTCTCTTT GATGTATTCG
151  AAGGAAACGG GCTGCCCGTG CCCTTCGTTG CGTAAAGATT CGTCCACGGG
201  CGGCAGGCCG ATGTCGCCGT GTATCCAACT TGCCAACCGC GATTGCGTGC
251  CGAAGGCGGA CACCTTGTTG CCTGTAACCG ACAGCACCAG CCCGCGTCCT
301  TTGCCTTTGG cggCTTCGCG CTTTTGGGCG AACAGCGCGT CAATCTGCGC
351  ATTCAATTCC GCCACGCGCG CTTCCTTACC GAAAATCCGC GACAGGGTCT
401  CCATCTGCTT CTCGCCGCTG GTGCGGATAT TGCCGTTGTC CACCGTCAAA
451  TCTATGgtgG TCGCGTTTTT CGCCAACTGT TCATACGCTT CCGCACCCGG
501  CCCGCCGGTA ATGACAAACT GCGGATTGTG GCGGTGCAGG GATTCGCAAT
551  CGGGCTCAAA CAGCGTCCCC ACCGTTGCCG CCTTGTCAAA TGCAGGCTGC
601  AAATAG
```

This corresponds to the amino acid sequence <SEQ ID 166; ORF 042.ng>:

g042.pep

```
  1  MTMICLRFQA FVPHTSALSN TSTAAGPSCP MAAVRSMMKI QPGFFSLMYS
 51  KETGCPCPSL RKDSSTGGRP MSPCIQLANR DCVPKADTLL PVTDSTSPRP
101  LPLAASRFWA NSASICAFNS ATRASLPKIR DRVSICFSPL VRILPLSTVK
151  SMVVAFFANC SYASAPGPPV MTNCG.WRCR DSQSGSNSVP TVAALSNAGC
201  K*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 167>:

m042.seq
```
  1  ATGACGATGA TTTGCTTGCG CTTCCAAGCG TTCGTGCCGC GTACCAGCGC
 51  GTTATCCAmT ACTTCGACAG CCGcCGGCCy TTCyTGCCCG ATGGCGGCGG
101  TACGGTCGAT GATGAAAATC CAATCGGGGT TTTTCTCTTT GATGTATTCG
151  AAGGAAACAG GCTGCCCGTG CCCCTCGTTG CGTAAAGATT CGTCTACAGG
201  CGGTAGGCCG ATGTCGCCGT GTATCCAACT TGCCAACCGC GACTGCGTGC
251  CGAAGGCGGA CACCTTGTTG CCCGTAACCG ACAGCACCAG CCCGCGTCCT
301  TTGCCTTTGG CGGCTTCGCG CGTTTGGGCG AACAGCGCGT CAATCTGCGC
351  CTTCAATTCC GCCGCGCGCG CTTCCTTGCC GAAAATCCGC GCCAAGGTCT
401  CCATCTGCTT TTCGCCGCTG GTGCGGATAT TGCCGTTGTC CACCGTCAGA
451  TCTATGGTGG TCGCGTTTTT CGCTAACTGT TCATACGCTT CCGCGCCCGG
501  CCCGCCGGTA ATGACAAGCT GAGGATTGTA GCGGTGCAGG GCTTCGTAAT
551  CGGGCTCGAA CAGCGTCCCC ACCGTTGCCG CCTTGTTAAA TGCAGGCTGC
601  AAATAA
```

This corresponds to the amino acid sequence <SEQ ID 168; ORF 042>:

m042.pep
```
  1  MTMICLRFQA FVPRTSALSX TSTAAGXSCP MAAVRSMMKI QSGFFSLMYS
 51  KETGCPCPSL RKDSSTGGRP MSPCIQLANR DCVPKADTLL PVTDSTSPRP
101  LPLAASRVWA NSASICAFNS AARASLPKIR AKVSICFSPL VRILPLSTVR
151  SMVVAFFANC SYASAPGPPV MTSXGLXRCR ASXSGSNSVP TVAALSNAGC
201  K*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 169>:

a042.seq
```
  1  ATGACGATGA TTTGCTTGCG CTTCCAAGCG TTCGTGCCGC GTACCAGCGC
 51  GTTATCCAAT ACTTCGACAG CCGCCGGCCC TTCCTGCCCG ATGGCGGCGG
101  TACGGTCGAT GATGAAAATC CAATCGGGGT TTTTCTCTTT GATGTATTCG
151  AAGGAAACAG GCTGCCCGTG CCCCTCGTTG CGTAAAGATT CGTCTACAGG
201  CGGTAGGCCG ATGTCGCCGT GTATCCAACT TGCCAACCGC GACTGCGTGC
251  CGAAGGCGGA CACCTTGTTG CCCGTAACCG ACAGCACCAG CCCGCGTCCT
301  TTGCCTTTGG CGGCTTCGCG CGTTTGGGCG AACAGCGCGT CAATCTGCGC
351  CTTCAATTCC GCCGCGCGCG CTTCCTTGCC GAAAATCCGC GCCAAGGTCT
401  CCATCTGCTT TTCGCCGCTG GTGCGGATAT TGCCGTTGTC CACCGTCAGA
451  TCTATGGTGG TCGCGTTTTT CGCCAACTGT TCATACGCTT CCGCGCCCGG
501  CCCGCCGGTA ATGACAAGCT GAGGATTGTA GCGGTGCAGG GCTTCGTAAT
551  CGGGCTCGAA CAGCGTCCCC ACCGTTGCCG CCTTGTCAAA TGCAGGCTGC
601  AAATAA
```

This corresponds to the amino acid sequence <SEQ ID 170; ORF 042.a>:

a042.pep
```
  1  MTMICLRFQA FVPRTSALSN TSTAAGPSCP MAAVRSMMKI QSGFFSLMYS
 51  KETGCPCPSL RKDSSTGGRP MSPCIQLANR DCVPKADTLL PVTDSTSPRP
101  LPLAASRVWA NSASICAFNS AARASLPKIR AKVSICFSPL VRILPLSTVR
151  SMVVAFFANC SYASAPGPPV MTS*GL*RCR AS*SGSNSVP TVAALSNAGC
201  K*
```

m042/a042 99.0% identity over a 201 aa overlap

```
                    10        20        30        40        50        60
m042.pep    MTMICLRFQAFVPRTSALSXTSTAAGXSCPMAAVRSMMKIQSGFFSLMYSKETGCPCPSL
            ||||||||||||||||||| |||||| |||||||||||||||||||||||||||||||||
a042        MTMICLRFQAFVPRTSALSNTSTAAGPSCPMAAVRSMMKIQSGFFSLMYSKETGCPCPSL
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m042.pep    RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVTDSTSPRPLPLAASRVWANSASICAFNS
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a042        RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVTDSTSPRPLPLAASRVWANSASICAFNS
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m042.pep    AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANCSYASAPGPPVMTSXGLXRCR
```

```
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a042         AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANCSYASAPGPPVMTSXGLXRCR
                     130       140       150       160       170       180

                     190       200
m042.pep     ASXSGSNSVPTVAALSNAGCKX
             ||||||||||||||||||||||
a042         ASXSGSNSVPTVAALSNAGCKX
                     190       200
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 042 shows 93.0% identity over a 201 aa overlap with a predicted ORF (ORF 042.ng) from *N. gonorrhoeae:*

m042/g042

```
                     10        20        30        40        50        60
m042.pep     MTMICLRFQAFVPRTSALSXTSTAAGXSCPMAAVRSMMKIQSGFFSLMYSKETGCPCPSL
             |||||||||||||:||||| |||||| ||||||||||||||| ||||||||||||||||||
g042         MTMICLRFQAFVPHTSALSNTSTAAGPSCPMAAVRSMMKIQPGFFSLMYSKETGCPCPSL
                     10        20        30        40        50        60

                     70        80        90       100       110       120
m042.pep     RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVTDSTSPRPLPLAASRVWANSASICAFNS
             |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
g042         RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVTDSTSPRPLPLAASRFWANSASICAFNS
                     70        80        90       100       110       120

                    130       140       150       160       170       180
m042.pep     AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANCSYASAPGPPVMTSXGLXRCR
             |:|||||||| :||||||||||||||||||:|||||||||||||||||||||||: || |||
g042         ATRASLPKIRDRVSICFSPLVRILPLSTVKSMVVAFFANCSYASAPGPPVMTNCGLWRCR
                    130       140       150       160       170       180

                    190       200
m042.pep     ASXSGSNSVPTVAALSNAGCKX
              | ||||||||||||||||||||
g042         DSQSGSNSVPTVAALSNAGCKX
                    190       200
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 171>:

```
m042-1.seq
      1  ATGACGATGA TTTGCTTGCG CTTCCAAGCG TTCGTGCCGC GTACCAGCGC
     51  GTTATCCAAT ACTTCGACAG CCGCCGGCCC TTCCTGCCCG ATGGCGGCGG
    101  TACGGTCGAT GATGAAAATC CAATCGGGGT TTTTCTCTTT GATGTATTCG
    151  AAGGAAACAG GCTGCCCGTG CCCCTCGTTG CGTAAAGATT CGTCTACAGG
    201  CGGTAGGCCG ATGTCGCCGT GTATCCAACT TGCCAACCGC GACTGCGTGC
    251  CGAAGGCGGA CACCTTGTTG CCCGTAACCG ACAGCACCAG CCCGCGTCCT
    301  TTGCCTTTGG CGGCTTCGCG CGTTTGGGCG AACAGCGCGT CAATCTGCGC
    351  CTTCAATTCC GCCGCGCGCG CTTCCTTGCC GAAAATCCGC GCCAAGGTCT
    401  CCATCTGCTT TTCGCCGCTG GTGCGGATAT TGCCGTTGTC CACCGTCAGA
    451  TCTATGGTGG TCGCGTTTTT CGCTAACTGT TCATACGCTT CCGCGCCCGG
    501  CCCGCCGGTA A
```

This corresponds to the amino acid sequence <SEQ ID 172; ORF 042-1>:

```
m042-1.pep
      1  MTMICLRFQA FVPRTSALSN TSTAAGPSCP MAAVRSMMKI QSGFFSLMYS
     51  KETGCPCPSL RKDSSTGGRP MSPCIQLANR DCVPKADTLL PVTDSTSPRP
    101  LPLAASRVWA NSASICAFNS AARASLPKIR AKVSICFSPL VRILPLSTVR
    151  SMVVAFFANC SYASAPGPPV MTS*
```

m042-1/g042 95.4% identity in 173 aa overlap

```
                    10        20        30        40        50        60
m042-1.pep  MTMICLRFQAFVPRTSALSNTSTAAGPSCPMAAVRSMMKIQSGFFSLMYSKETGCPCPSL
            ||||||||||||:||||||||||||||||||||||||||||||| |||||||||||||||
g042        MTMICLRFQAFVPHTSALSNTSTAAGPSCPMAAVRSMMKIQPGFFSLMYSKETGCPCPSL
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m042-1.pep  RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVTDSTSPRPLPLAASRVWANSASICAFNS
            ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
g042        RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVTDSTSPRPLPLAASRFWANSASICAFNS
                    70        80        90       100       110       120

                   130       140       150       160       170
m042-1.pep  AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANCSYASAPGPPVMTSX
            |:|||||||| :|||||||||||||||||:||||||||||||||||||||||:
g042        ATRASLPKIRDRVSICFSPLVRILPLSTVKSMVVAFFANCSYASAPGPPVMTNCGLWRCR
                   130       140       150       160       170       180

g042        DSQSGSNSVPTVAALSNAGCKX
                   190       200
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 173>:

```
a042-1.seq
      1  ATGACGATGA TTTGCTTGCG CTTCCAAGCG TTCGTGCCGC GTACCAGCGC
     51  GTTATCCAAT ACTTCGACAG CCGCCGGCCC TTCCTGCCCG ATGGCGGCGG
    101  TACGGTCGAT GATGAAAATC CAATCGGGGT TTTTCTCTTT GATGTATTCG
    151  AAGGAAACAG GCTGCCCGTG CCCCTCGTTG CGTAAAGATT CGTCTACAGG
    201  CGGTAGGCCG ATGTCGCCGT GTATCCAACT TGCCAACCGC GACTGCGTGC
    251  CGAAGGCGGA CACCTTGTTG CCCGTAACCG ACAGCACCAG CCCGCGTCCT
    301  TTGCCTTTGG CGGCTTCGCG CGTTTGGGCG AACAGCGCGT CAATCTGCGC
    351  CTTCAATTCC GCCGCGCGCG CTTCCTTGCC GAAAATCCGC GCCAAGGTCT
    401  CCATCTGCTT TTCGCCGCTG GTGCGGATAT TGCCGTTGTC CACCGTCAGA
    451  TCTATGGTGG TCGCGTTTTT CGCCAACTGT TCATACGCTT CCGCGCCCGG
    501  CCCGCCGGTA A
```

This corresponds to the amino acid sequence <SEQ ID 174; ORF 042-1.a>:

```
a042-1.pep
      1  MTMICLRFQA FVPRTSALSN TSTAAGPSCP MAAVRSMMKI QSGFFSLMYS
     51  KETGCPCPSL RKDSSTGGRP MSPCIQLANR DCVPKADTLL PVTDSTSPRP
    101  LPLAASRVWA NSASICAFNS AARASLPKIR AKVSICFSPL VRILPLSTVR
    151  SMVVAFFANC SYASAPGPPV MTS*
```

m042-1/a042-1   100.0% identity in 173 aa overlap

```
                    10        20        30        40        50        60
m042-1.pep  MTMICLRFQAFVPRTSALSNTSTAAGPSCPMAAVRSMMKIQSGFFSLMYSKETGCPCPSL
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a042-1      MTMICLRFQAFVPRTSALSNTSTAAGPSCPMAAVRSMMKIQSGFFSLMYSKETGCPCPSL
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m042-1.pep  RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVTDSTSPRPLPLAASRVWANSASICAFNS
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a042-1      RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVTDSTSPRPLPLAASRVWANSASICAFNS
                    70        80        90       100       110       120

                   130       140       150       160       170
m042-1.pep  AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANCSYASAPGPPVMTSX
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
a042-1      AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANCSYASAPGPPVMTSX
                   130       140       150       160       170
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 175>:

```
g043.seq
     1  ATGGTTGTTT CAAATCAAAA TATCTATGCC GTCGGCCCAT CAGCACTTTT
    51  TCACATCCGA AGGCAAAAAT CCGTAATGCC GCCTGAACGC TTCgttgaAC
   101  CGTCCCGCGT ggcggtagcc gcAAAAGTGC ATcGCGGCTT GGATGGTGCT
   151  GCCCGATTCG ATGAGGGcga gcGCGTGTTC CAGCCGCAGG CGGCGCAGGC
   201  GTCCGGCGAC GGTTTCGCCG GTTTGCGCTT TGAAATAGCG TTTCAGGTAG
   251  CATTCGTTCA GCCCGACGCG GCGGGCGATT TCGGCGATGG TCAGCGGGCG
   301  GGCGAATTCG CTGTTCAAAA TATCGGCGGC TTCGTCTATG CGCCGGCGGC
   351  GGTAGCCGTT GTCGTGGCGG CGGAAGGTGA AGCGTAA
```

This corresponds to the amino acid sequence <SEQ ID 176; ORF 043.ng>:

```
g043.pep
     1  MVVSNQNIYA VGPSALFHIR RQKSVMPPER FVEPSRVAVA AKVHRGLDGA
    51  ARFDEGERVF QPQAAQASGD GFAGLRFEIA FQVAFVQPDA AGDFGDGQRA
   101  GEFAVQNIGG FVYAPAAVAV VVAAEGEA*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 177>:

```
m043.seq
     1  ATGGTTGTTT CAAATCAAAA TATCTATGCC GCCGGCCCCT CAGCACTTCT
    51  TCACATCCGA AGGCAAAAAT CCGTAATGCC GTCTGAACGC TTCGTTGAAC
   101  CGTCCCGCGT GGCGGTAGCC GCAAAAGTGC ATGGCGGCTT GGACGGTGCT
   151  GCCGGATTCG ATGAGGGCGA GCGCGTGTTC CAGCCGCAGG CGGCGCAgGC
   201  ATCCGGCGAC GGTTTCGCCG GTTTGCGCTT TGAAATAGCG TTTCAGGTAG
   251  CATTCGTTCA GTCCGACGCG GCGGGCGATT TCGGCGATGG TCAGCGGACG
   301  GGCGAATTCG TGTTGCAGGA TGTCGGCGGC TTCGTCTATG CGCCGACGGC
   351  GGTAACCGTT GTCGTGGCGG CGGAAGGTGA AGCGCAATAA
```

This corresponds to the amino acid sequence <SEQ ID 178; ORF 043>:

```
m043.pep
     1  MVVSNQNIYA AGPSALLHIR RQKSVMPSER FVEPSRVAVA AKVHGGLDGA
    51  AGFDEGERVF QPQAAQASGD GFAGLRFEIA FQVAFVQSDA AGDFGDGQRT
   101  GEFVLQDVGG FVYAPTAVTV VVAAEGEAQ*
```

Computer analysis of the amino acid sequences gave the following results:

Homology with a predicted ORF from *N. meningitidis menA with menB*

ORF 043 shows 89.8% identity over a 128 aa overlap with a predicted ORF (ORF043.a) from *N. gonorrhoeae:*

m043/g043

```
                     10        20        30        40        50        60
m043.pep     MVVSNQNIYAAGPSALLHIRRQKSVMPSERFVEPSRVAVAAKVHGGLDGAAGFDEGERVF
             ||||||||||:|||||:||||||||||| |||||||||||||||| |||||| |||||||
g043         MVVSNQNIYAVGPSALFHIRRQKSVMPPERFVEPSRVAVAAKVHRGLDGAARFDEGERVF
                     10        20        30        40        50        60

                     70        80        90       100       110       120
m043.pep     QPQAAQASGDGFAGLRFEIAFQVAFVQSDAAGDFGDGQRTGEFVLQDVGGFVYAPTAVTV
             ||||||||||||||||||||||||||| |||||||||||:|||::|::|||||||:||:|
g043         QPQAAQASGDGFAGLRFEIAFQVAFVQPDAAGDFGDGQRAGEFAVQNIGGFVYAPAAVAV
                     70        80        90       100       110       120

                    130
m043.pep     VVAAEGEAQX
             ||||||||
g043         VVAAEGEAXX
                    130
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 179>:

```
a043.seq
     1  ATGGTTGTTT CAAATCAAAA TATCTATGCC GCCGGCCCCT CAGCACTTCT
    51  TCACATCCGA AGGCAAAAAT CCGTAATGCC GTCTGAACGC TTCGTTGAAC
   101  CGTCCCGCGT GGCGGTAGCC GCAAAAGTGC ATGGCGGCTT GGACGGTGCT
```

```
151   GCCGGATTCG ATGAGGGCGA GCGCGTGTTC CAGCCGCAGG CGGCGCAGGC
201   ATCCGGCGAC GGTTTCGCCG GTTTGCGCTT TGAAATAGCG TTTCAGGTAG
251   CATTCGTTCA GTCCGACGCG GCGGGCGATT TCGGCGATGG TCAGCGGACG
301   GGCGAATTCG TGTTGCAGGA TGTCGGCGGC TTCGTCTATG CGCCGACGGC
351   GGTAACCGTT GTCGTGGCGG CGGAAGGTGA AGCGCAATAA
```

This corresponds to the amino acid sequence <SEQ ID 180; ORF 043.a>:

```
a043.pep
  1   MVVSNQNIYA AGPSALLHIR RQKSVMPSER FVEPSRVAVA AKVHGGLDGA
 51   AGFDEGERVF QPQAAQASGD GFAGLRFEIA FQVAFVQSDA AGDFGDGQRT
101   GEFVLQDVGG FVYAPTAVTV VVAAEGEAQ*
```

m043/a043 100.0% identity in 129 aa overlap

```
                    10        20        30        40        50        60
m043.pep    MVVSNQNIYAAGPSALLHIRRQKSVMPSERFVEPSRVAVAAKVHGGLDGAAGFDEGERVF
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a043        MVVSNQNIYAAGPSALLHIRRQKSVMPSERFVEPSRVAVAAKVHGGLDGAAGFDEGERVF
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m043.pep    QPQAAQASGDGFAGLRFEIAFQVAFVQSDAAGDFGDGQRTGEFVLQDVGGFVYAPTAVTV
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a043        QPQAAQASGDGFAGLRFEIAFQVAFVQSDAAGDFGDGQRTGEFVLQDVGGFVYAPTAVTV
                    70        80        90       100       110       120

                   130
m043.pep    VVAAEGEAQX
            ||||||||||
a043        VVAAEGEAQX
                   130
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 181>:

```
g044.seq
  1   ATGCTGCCCG ACCAGAGCGT CGAGTTCTTG CCACAAGTCG TCGTTTTTGA
 51   CGGGCTGTTT GGCGGCGGTT TTCCAGCCGT TGCGCTTCCA ACCGTGTATC
101   CAGTTTTCCA TGCCGTTTTT GACGTATTGC GAGTCGGTGC AGATGATGAC
151   GGTGCAGCGG CGTTTGAGCG ATTTCAGCCC TTCGATAACG GCGGTCAGCT
201   CCATGCGGTT GTTGGTGGTT TGCGCTTCGC CGCCGAAAAG TTCTTTTTCG
251   CGGCTGCCGT AGCGCATTAA
```

This corresponds to the amino acid sequence <SEQ ID 182; ORF 044.ng>:

```
g044.pep
  1   MLPDQSVEFL PQVVVFDGLF GGGFPAVALP TVYPVFHAVF DVLRVGADDD
 51   GAAAFERFQP FDNGGQLHAV VGGLRFAAEK FFFAAAVAH*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 183>:

```
m044.seq
  1   ATGCCGTCCG ACTAGAGCGT CGAGTTCTTT CCAGAAGTCG TCGTTTTTGA
 51   CGGGCTGTTT GGAGGCGGTT TTCCAGCCGT TGCGCTTCCA ACCGTGTATC
101   CAGTTTTCCA TGCCATTTTT GACGTATTGC GAGTCGGTGC AGATGATGAC
151   GGTGCAGCGG CGTTTGAGCG ATTTCAGTCC TTCGATGACG GCAGTCAGTT
201   CCATGCGGTT GTTGGTGGTT TGCGCTTCGC CGCCGAAAAG TTCTTTTTCG
251   TGGCTACCGT AGCGCAyTAa
```

This corresponds to the amino acid sequence <SEQ ID 184; ORF 044>:

```
m044.pep
  1   MPSDXSVEFF PEVVVFDGLF GGGFPAVALP TVYPVFHAIF DVLRVGADDD
 51   GAAAFERFQS FDDGSQFHAV VGGLRFAAEK FFFVATVAH*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 185>:

```
a044.seq
      1  GTGCCGTCCG ACCAGCGCGT CGAGTTCTTT CCACAAGTCG TCGTTTTTGA
     51  CGGGCTGTTT GGCGGCGGTT TTCCAGCCGT TGCGCTTCCA ACCGTGTATC
    101  CAGTTTTCCA TGCCGTTTTT GACGTATTGC GAGTCGGTGC AGATGATGAC
    151  GGTGCAGCGG CGTTTGAGCG ATTTCAGTCC TTCGATGACG GCGGTCAGTT
    201  CCATACGGTT GTTGGTGGTT TGCGCTTCGC CGCCGAAAAG TTCTTTTTCG
    251  TGGCTGCCGT AGCGCATTAA
```

This corresponds to the amino acid sequence <SEQ ID 186; ORF 044.a>:

```
a044.pep
      1  VPSDQRVEFF PQVVVFDGLF GGGFPAVALP TVYPVFHAVF DVLRVGADDD
     51  GAAAFERFQS FDDGGQFHTV VGGLRFAAEK FFFVAAVAH*
```

m044/a044 91.0% identity over a 89 aa overlap

```
                        10        20        30        40        50        60
m044.pep        MPSDXSVEFFPEVVVFDGLFGGGFPAVALPTVYPVFHAIFDVLRVGADDDGAAAFERFQS
                :|||  |||||:|||||||||||||||||||||||||||||||:||||||||||||||||
a044            VPSDQRVEFFPQVVVFDGLFGGGFPAVALPTVYPVFHAVFDVLRVGADDDGAAAFERFQS
                        10        20        30        40        50        60

                        70        80        90
m044.pep        FDDGSQFHAVVGGLRFAAEKFFFVATVAHX
                ||||:|||:||||||||||||||||:||||
a044            FDDGGQFHTVVGGLRFAAEKFFFVAAVAHX
                        70        80        90
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 044 shows 86.5% identity over a 89 aa overlap with a predicted ORF (ORF 044.ng) from *N. gonorrhoeae:*

m044/g044

```
                        10        20        30        40        50        60
m044.pep        MPSDXSVEFFPEVVVFDGLFGGGFPAVALPTVYPVFHAIFDVLRVGADDDGAAAFERFQS
                |  | ||||:|:|||||||||||||||||||||||||||||||:||||||||||||||| 
g044            MLPDQSVEFLPQVVVFDGLFGGGFPAVALPTVYPVFHAVFDVLRVGADDDGAAAFERFQP
                        10        20        30        40        50        60

                        70        80        90
m044.pep        FDDGSQFHAVVGGLRFAAEKFFFVATVAHX
                ||:|:|:||||||||||||||||:|:||||
g044            FDNGGQLHAVVGGLRFAAEKFFFAAAVAHX
                        70        80        90
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 187>:

```
g046.seq
      1  ATGTCGGCAA TGCTGCGTCC GACAAGCAGC CCGCCGCgcc gCGCCTGTAT
     51  GATGACCATC CGCACGCGGT CGTCTGCAAA ACGTAAAACC TGCAATGCGC
    101  CCGGGCAGTC TATCAGGCCG GCAAGCTGTT CGGTAACGAG CTGTTCGGGG
    151  CTGATGGTTT CGGTTATGCC gaATATGGAA AGGCTGCCGt TTTcGTTGTT
    201  TTCGAGCTTG GGGCTGAGGT ATTCGAGGTA TtcgctGGAA CGGACGCGCG
    251  CGATGCGGCC GGGGATGTTG AACAGGTCGG CGGCAACTTT GCAGGCGACG
    301  ATGTTGGTTT CGTCGCTGCG GGagaGCGCG AGcagcaagt cggcatcttC
    351  CgcgccggcG Cgttataatg tgAAGGGGGA TGCGccgttg ccgaAAACGG
    401  TTTGGacatc gaggcggctg CCTGTTTCCT GCAATGCTTT TTCGTCGATG
    451  TCGATAAcgg TTACGTCCTT GTTGGTGATG GCGGCAAGGT TTTGCGCGAC
```

```
501  GGTAGAACCT ACCTGCCCGT TGCCTAAAAT GAGGATTTTC ACGGTATGGG
551  TCGCCGGGTG A
```

This corresponds to the amino acid sequence <SEQ ID 188; ORF 046.ng>:

```
g046.pep
  1  MSAMLRPTSS PPRRACMMTI RTRSSAKRKT CNAPGQSIRP ASCSVTSCSG
 51  LMVSVMPNME RLPFSLFSSL GLRYSRYSLE RTRAMRPGML NRSAATLQAT
101  MLVSSLRESA SSKSASSAPA RYNVKGDAPL PKTVWTSRRL PVSCNAFSSM
151  SITVTSLLVM AARFCATVEP TCPLPKMRIF TVWVAG*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 189>:

```
m046.seq
  1  ATGTCGGCAA TGCTGCGTCC GACAAGCAsT CCGC.r.sGC gCGcCTGTAT
 51  GATGACCATC CGCACGCGGT CGTCTGCAAA ACGTAAAACC TGCAATGCGC
101  CCGGGCAGTC TATCAGGCCG GCAAGCTGTT CGGTAACGAG CTGTTCGGGG
151  CTGATGGTTT CGGTTATGCC GAATATGGAA AGGCTGCCGT TTTCGTTGTT
201  TTCGAGCTTG GGGCTGAGGT ATTCGAGGTA TTCGCTGGAA CGGACGCGCG
251  CGATGCGGCC GGGGATGTTG AACAGGTCGG CGGCAACTTT GCAGGCGACG
301  ATGTTGGTTT CGTCGCTGCG GGAGAGCGCG AGCAGCAAGT CGGCATCTTC
351  CGCGCCGGCG CGTTCTAATG TGAAGGGGGA TGCGCCGTTG CCGAAAACGG
401  TTTGGACATC GAGGCGGCTG CCTGTTTCCT GCAATGCTTT TTCGTCGATG
451  TCGATAACGG TTACGTCGTT GTTGGGTATG GCGGCAAGGT TTTGTGCGAC
501  GGTAGAACCT ACCTGTCCGT TGCCTAAAAT GAGGATTTTC ACGGTGTGGG
551  TCGCCGAGTG A
```

This corresponds to the amino acid sequence <SEQ ID 190; ORF 046>:

```
m046.pep
  1  MSAMLRPTSX PXXRACMMTI RTRSSAKRKT CNAPGQSIRP ASCSVTSCSG
 51  LMVSVMPNME RLPFSLFSSL GLRYSRYSLE RTRAMRPGML NRSAATLQAT
101  MLVSSLRESA SSKSASSAPA RSNVKGDAPL PKTVWTSRRL PVSCNAFSSM
151  SITVTSLLGM AARFCATVEP TCPLPKMRIF TVWVAE*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 191>:

```
a046.seq
  1  ATGTCGGCAA TGCTGCGTCC GACAAGCAGT CCGCCGCGCC GCGCCTGTAT
 51  GATGACCATC CGCACGCGGT CGTCTGCAAA ACGTAAAACC TGCAATGCGC
101  CCGGGCAGTC TATCAGGCCG GCAAGCTGTT CGGTAACGAG CTGTTCGGGG
151  CTGATGGTTT CGGTTATGCC GAATATGGAA AGGCTGCCGT TTTCGTTGTT
201  TTCGAGCTTG GGGCTGAGGT ATTCGAGGTA TTCGCTGGAA CGGACGCGCG
251  CGATGCGGCC GGGGATGTTG AACAGGTCGG CGGCAACTTT GCAGGCGACG
301  ATGTTGGTTT CGTCGCTGCG GGAGAGCGCG AGCAGCAAGT CGGCATCTTC
351  CGCGCCGGCG CGTTCTAATG TGAAGGGGGA TGCGCCGTTG CCGAAAACGG
401  TTTGGACATC GAGGCGGCTG CCTGTTTCCT GCAATGCTTT TTCGTCGATG
451  TCGATAACGG TTACGTCGTT GTTGGGTATG GCGGCAAGGT TTTGTGCGAC
501  GGTAGAACCT ACCTGTCCGT TGCCTAAAAT GAGGATTTTC ACGGTGTGGG
551  TCGCCGAGTG A
```

This corresponds to the amino acid sequence <SEQ ID 192; ORF 046.a>:

```
a046.pep
  1  MSAMLRPTSS PPRRACMMTI RTRSSAKRKT CNAPGQSIRP ASCSVTSCSG
 51  LMVSVMPNME RLPFSLFSSL GLRYSRYSLE RTRAMRPGML NRSAATLQAT
101  MLVSSLRESA SSKSASSAPA RSNVKGDAPL PKTVWTSRRL PVSCNAFSSM
151  SITVTSLLGM AARFCATVEP TCPLPKMRIF TVWVAE*
```

m046/a046 98.4% identity over a 186 aa overlap

```
                   10        20        30        40        50        60
m046.pep    MSAMLRPTSXPXXRACMMTIRTRSSAKRKTCNAPGQSIRPASCSVTSCSGLMVSVMPNME
            ||||||||| |  |||||||||||||||||||||||||||||||||||||||||||||||
a046        MSAMLRPTSSPPRRACMMTIRTRSSAKRKTCNAPGQSIRPASCSVTSCSGLMVSVMPNME
```

```
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m046.pep    RLPFSLFSSLGLRYSRYSLERTRAMRPGMLNRSAATLQATMLVSSLRESASSKSASSAPA
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a046        RLPFSLFSSLGLRYSRYSLERTRAMRPGMLNRSAATLQATMLVSSLRESASSKSASSAPA
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m046.pep    RSNVKGDAPLPKTVWTSRRLPVSCNAFSSMSITVTSLLGMAARFCATVEPTCPLPKMRIF
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a046        RSNVKGDAPLPKTVWTSRRLPVSCNAFSSMSITVTSLLGMAARFCATVEPTCPLPKMRIF
                   130       140       150       160       170       180

m046.pep    TVWVAEX
            |||||||
a046        TVWVAEX
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

ORF 046 shows 97.3% identity over a 185 aa overlap with a predicted ORF (ORF 046.ng) from *N. gonorrhoeae:*

```
m046/g046

                    10        20        30        40        50        60
m046.pep    MSAMLRPTSXPXXRACMMTIRTRSSAKRKTCNAPGQSIRPASCSVTSCSGLMVSVMPNME
            ||||||||| |  |||||||||||||||||||||||||||||||||||||||||||||||
g046        MSAMLRPTSSPPRRACMMTIRTRSSAKRKTCNAPGQSIRPASCSVTSCSGLMVSVMPNME
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m046.pep    RLPFSLFSSLGLRYSRYSLERTRAMRPGMLNRSAATLQATMLVSSLRESASSKSASSAPA
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g046        RLPFSLFSSLGLRYSRYSLERTRAMRPGMLNRSAATLQATMLVSSLRESASSKSASSAPA
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m046.pep    RSNVKGDAPLPKTVWTSRRLPVSCNAFSSMSITVTSLLGMAARFCATVEPTCPLPKMRIF
            | |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
g046        RYNVKGDAPLPKTVWTSRRLPVSCNAFSSMSITVTSLLVMAARFCATVEPTCPLPKMRIF
                   130       140       150       160       170       180

m046.pep    TVWVAEX
            |||||
g046        TVWVAGX
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 193>:

```
g047.seq
      1  ATGGTCATCA TACAGGCGcg gcGCGGCGGG CTGCTTGTCG GACGCAGCAT
     51  TGCCGACATC GCCCAAGATT TGCCCGACGG GGCCGACTGC CAAATCTGCG
    101  CCGTTTACCG CAACAACCGC CTCATC[illegible]TCC CCGCGCCGCA AACCGTCATC
    151  ATCGAAGGCG ACGAAATCCT GTTTGCCGCC GCCGCCGAAA ACATCGGGGC
    201  GGTCATACCc gaATTGCGCC CCAAAGAAAC CAGCACCCGC CGCATCATGA
    251  TTGCCGGCGG CGGCAACATc tgctACCGCC TCGCCAAGCA GCTCGAACAC
```

301 GCATAcaacG TCAAAATCAT CGAATGCCGG CCGCGCcgtg ccgaATGGAT
351 AGCCGAAAAC ctcgAcaaCA CCCTCGTCCT GCAAGGTTCG Gcaaccgacg
401 aAaccctgct cgAcaacgaa tacatcgacg aaatcgaCGT ATTCTGCGCC
451 CTGACCAACG ACGACGAAAG CAACATTAtg tCCGCCCTTT TGGCGAAAAA
501 CCTcggcgCG AAGCgcgtca tcggCATCGT CAACCGCTCA AGCTACGTCG
551 ATTTGCTCGA AGGCAACAAA ATCGACATCG TCGTCTCCCC CCACCTCATC
601 ACCATCGGCT CGATACTCGC CCACATCCGG CGCGGCGACA TCGTTGCCGT
651 CCACCCCATC CGGCGCGGCA CGGCGGAAGC CATCGAAGTC GTCGCGCACG
701 GCGACAAAAA AACTTCCGCC ATCATCGGCA GGCGCATCAG CGGCATCAAA
751 TGGCCCGAAG GCTGCCACAT TGCCGCCGTC GTCCGCGCCG GAACCGGCGA
801 AACCATTATG GGACACCATA CCGAAACCGT CATCCAAGAC GGTGACCACA
851 TCATCTTTTT CGTCTCGCGC CGGCGCATCC TGAACGAACT GGAGAAACTC
901 ATCCAAGTCA AAATGGGCTT TTTCGGATAA

This corresponds to the amino acid sequence <SEQ ID 194; ORF 047.ng>:

g047.pep

1 MVIIQARRGG LLVGRSIADI AQDLPDGADC QICAVYRNNR LIVPAPQTVI
51 IEGDEILFAA AAENIGAVIP ELRPKETSTR RIMIAGGGNI CYRLAKQLEH
101 AYNVKIIECR PRRAEWIAEN LDNTLVLQGS ATDETLLDNE YIDEIDVFCA
151 LTNDDESNIM SALLAKNLGA KRVIGIVNRS SYVDLLEGNK IDIVVSPHLI
201 TIGSILAHIR RGDIVAVHPI RRGTAEAIEV VAHGDKKTSA IIGRRISGIK
251 WPEGCHIAAV VRAGTGETIM GHHTETVIQD GDHIIFFVSR RRILNELEKL
301 IQVKMGFFG*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 195>:

m047.seq

1 ATGGTCATCA TACAGgCGcG C..syGCGGA sTGCTTGTCG GACGCAGCAT
51 TGCCGACATC GCCCAAGATT TGCCCGACGG GGCCGACTGC CAAATCTGCG
101 CCGTTTACCG CAACAACCGC CTCATCGTCC CCGCGCCGCA AACCGTCATC
151 ATCGAAGGCG ACGAAATCCT ATTTGCCGCC GCCGCCGAAA ACATCGGCGC
201 GGTCATACCC GAATTGCGCC CCAAAGAAAC CCAAAGAAAC CAGCcCmgmm
251 GcATCATGAT TkCCGGCGGC GGCAACATCG GCTACCGTCT CGCCAAGCAG
301 CTCGAACACG CATACAACGT yAAAATCATC GAATGCCGGC CGCGCCGTGC
351 CGAATGGATA GCCGAAAACC TCGACAACAC CCTCGTCyTG CAAGGTTCGG
401 CAACCGACGA AACCCTGCTC GACAACGAAT ACATCGACGA AATCGACGTA
451 TTCTGCGCCC TGACCAACGA CGACGAAAGC AACATTATGT CCGCCCTTTT
501 GGCGAaAAAC CTCGGCGCGA AGCGCGTCAT CGGCATCGTC AACCGCTCAA
551 GCTACGTCGA TTTGCTCGAA GGCAACAAAA TCGACATCGT CGTCTCCCCC
601 CACCTCATCA CCATCGGCTC GATACTCGCC CACATCCGGC GCGGCGACAT
651 CGTTGCCGTC CACCCCATCC GGCGCGGCAC GGCGGAAGCC ATCGAAGTCG
701 TCGCACACGG CGACAAAAAA ACTTCCGCCA TCATCGGCAG GCGCATCAGC
751 GGCATCAAAT GGCCCGAAGG CTGCCACATT GCCGCCGTCG TCCGCGCCGG
801 AACCGGCGAA ACCATTATGG GACACCATAC CGAAACCGTC ATCCAAGACG
851 GCGACCACAT CATCTTTTTC GTCTCGCGCC GGCGCATCCT GAACGAACTG
901 GAAAAACTCA TCCAGGTCAA AATGGGCTTT TTCGGATAA

This corresponds to the amino acid sequence <SEQ ID 196; ORF 047>:

m047.pep

1 MVIIQARXXG XLVGRSIADI AQDLPDGADC QICAVYRNNR LIVPAPQTVI
51 IEGDEILFAA AAENIGAVIP ELRPKETQRN QPXXIMIXGG GNIGYRLAKQ
101 LEHAYNVKII ECRPRRAEWI AENLDNTLVL QGSATDETLL DNEYIDEIDV
151 FCALTNDDES NIMSALLAKN LGAKRVIGIV NRSSYVDLLE GNKIDIVVSP
201 HLITIGSILA HIRRGDIVAV HPIRRGTAEA IEVVAHGDKK TSAIIGRRIS
251 GIKWPEGCHI AAVVRAGTGE TIMGHHTETV IQDGDHIIFF VSRRRILNEL
301 EKLIQVKMGF FG*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 197>:

a047.seq

1 ATGGTCATCA TACAGGCGCG GCGCGGCGGA CTGCTTGTCG GACGCAGCAT

```
 51  TGCCGACATC GCCCAAGATT TGCCCGACGG GGCCGACTGC CAAATCTGCG
101  CCGTTTACCG CAACAACCGC CTCATCGTCC CCGCGCCGCA AACCGTCATC
151  ATCGAAGGCG ACGAAATCCT ATTTGCCGCC GCCGCCGAAA ACATCGGCGC
201  GGTCATACCC GAATTGCGCC CCAAAGAAAC CAGCACCCGC CGCATCATGA
251  TTGCCGGCGG CGGCAACATC GGCTACCGTC TCGCCAAGCA GCTCGAACAC
301  GCATACAACG TCAAAATCAT CGAATGCCGG CCGCGCCGTG CCGAATGGAT
351  AGCCGAAAAC CTCGACAACA CCCTCGTCCT GCAAGGTTCG GCAACCGACG
401  AAACCCTGCT CGACAACGAA TACATCGACG AAATCGACGT ATTCTGCGCC
451  CTGACCAACG ACGACGAAAG CAACATTATG TCCGCCCTTT TGGCGAAAAA
501  CCTCGGCGCG AAGCGCGTCA TCGGCATCGT CAACCGCTCA AGCTACGTCG
551  ATTTGCTCGA AGGCAACAAA ATCGACATCG TCGTCTCCCC CCACCTCATC
601  ACCATCGGCT CGATACTCGC CCACATCCGG CGCGGCGACA TCGTTGCCGT
651  CCACCCCATC CGGCGCGGCA CGGCGGAAGC CATCGAAGTC GTCGCACACG
701  GCGACAAAAA AACTTCCGCC ATCATCGGCA GGCGCATCAG CGGCATCAAA
751  TGGCCCGAAG GCTGCCACAT TGCCGCCGTC GTCCGCGCCG GAACCGGCGA
801  AACCATTATG GGACACCATA CCGAAACCGT CATCCAAGAC GGCGACCACA
851  TCATCTTTTT CGTCTCGCGC CGGCGCATCC TGAACGAACT GGAAAAACTC
901  ATCCAAGTCA AAATGGGCTT TTTCGGATAA
```

This corresponds to the amino acid sequence <SEQ ID 198; ORF 047.a>:

```
a047.pep
  1  MVIIQARRGG LLVGRSIADI AQDLPDGADC QICAVYRNNR LIVPAPQTVI
 51  IEGDEILFAA AAENIGAVIP ELRPKETSTR RIMIAGGGNI GYRLAKQLEH
101  AYNVKIIECR PRRAEWIAEN LDNTLVLQGS ATDETLLDNE YIDEIDVFCA
151  LTNDDESNIM SALLAKNLGA KRVIGIVNRS SYVDLLEGNK IDIVVSPHLI
201  TIGSILAHIR RGDIVAVHPI RRGTAEAIEV VAHGDKKTSA IIGRRISGIK
251  WPEGCHIAAV VRAGTGETIM GHHTETVIQD GDHIIFFVSR RRILNELEKL
301  IQVKMGFFG*
```

m047/a047 96.5% identity over a 312 aa overlap

```
                   10        20        30        40        50        60
m047.pep   MVIIQARXXGXLVGRSIADIAQDLPDGADCQICAVYRNNRLIVPAPQTVIIEGDEILFAA
           |||||||  | |||||||||||||||||||||||||||||||||||||||||||||||||
a047       MVIIQARRGGLLVGRSIADIAQDLPDGADCQICAVYRNNRLIVPAPQTVIIEGDEILFAA
                   10        20        30        40        50        60

                   70        80        90       100       110       120
m047.pep   AAENIGAVIPELRPKETQRNQPXXIMIXGGGNIGYRLAKQLEHAYNVKIIECRPRRAEWI
           ||||||||||||||||:  :    ||| ||||||||||||||||||||||||||||||||
a047       AAENIGAVIPELRPKETSTRR---IMIAGGGNIGYRLAKQLEHAYNVKIIECRPRRAEWI
                   70        80           90       100       110

                  130       140       150       160       170       180
m047.pep   AENLDNTLVLQGSATDETLLDNEYIDEIDVFCALTNDDESNIMSALLAKNLGAKRVIGIV
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a047       AENLDNTLVLQGSATDETLLDNEYIDEIDVFCALTNDDESNIMSALLAKNLGAKRVIGIV
           120       130       140       150       160       170

                  190       200       210       220       230       240
m047.pep   NRSSYVDLLEGNKIDIVVSPHLITIGSILAHIRRGDIVAVHPIRRGTAEAIEVVAHGDKK
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a047       NRSSYVDLLEGNKIDIVVSPHLITIGSILAHIRRGDIVAVHPIRRGTAEAIEVVAHGDKK
           180       190       200       210       220       230

                  250       260       270       280       290       300
m047.pep   TSAIIGRRISGIKWPEGCHIAAVVRAGTGETIMGHHTETVIQDGDHIIFFVSRRRILNEL
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a047       TSAIIGRRISGIKWPEGCHIAAVVRAGTGETIMGHHTETVIQDGDHIIFFVSRRRILNEL
           240       250       260       270       280       290

                  310
m047.pep   EKLIQVKMGFFGX
           ||||||||||||
```

```
a047              EKLIQVKMGFFGX
                  300       310
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 047 shows 96.2% identity over a 312 aa overlap with a predicted ORF (ORF 047.ng) from *N. gonorrhoeae:*

m047/g045

```
m047.pep          MVIIQARXXGXLVGRSIADIAQDLPDGADCQICAVYRNNRLIVPAPQTVIIEGDEILFAA    60
                  |||||||  | |||||||||||||||||||||||||||||||||||||||||||||||||
g047              MVIIQARRGGLLVGRSIADIAQDLPDGADCQICAVYRNNRLIVPAPQTVIIEGDEILFAA    60

m047.pep          AAENIGAVIPELRPKETQRNQPXXIMIXGGGNIGYRLAKQLEHAYNVKIIECRPRRAEWI   120
                  |||||||||||||||||:  :   ||| ||||| ||||||||||||||||||||||||||
g047              AAENIGAVIPELRPKETSTRR---IMIAGGGNICYRLAKQLEHAYNVKIIECRPRRAEWI   117

m047.pep          AENLDNTLVLQGSATDETLLDNEYIDEIDVFCALTNDDESNIMSALLAKNLGAKRVIGIV   180
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g047              AENLDNTLVLQGSATDETLLDNEYIDEIDVFCALTNDDESNIMSALLAKNLGAKRVIGIV   177

m047.pep          NRSSYVDLLEGNKIDIVVSPHLITIGSILAHIRRGDIVAVHPIRRGTAEAIEVVAHGDKK   240
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g047              NRSSYVDLLEGNKIDIVVSPHLITIGSILAHIRRGDIVAVHPIRRGTAEAIEVVAHGDKK   237

m047.pep          TSAIIGRRISGIKWPEGCHIAAVVRAGTGETIMGHHTETVIQDGDHIIFFVSRRRILNEL   300
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g047              TSAIIGRRISGIKWPEGCHIAAVVRAGTGETIMGHHTETVIQDGDHIIFFVSRRRILNEL   297

m047.pep          EKLIQVKMGFFGX  313
                  ||||||||||||
g047              EKLIQVKMGFFGX  310
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 199>:

```
g048.seq
      1  ATGCTCGACA AAGGCGAGGA GTTGCCCGTC GATTTCACCA ACCGCCTGAT
     51  TTACTACGTc ggcCCcgTCG ATCCGGTCGG CGATGAAGTC GTCGGTCCCG
    101  CAGGTCCGAC CACAGCCACC CGCATGGACA AATTTACCCG CCAAATGCTC
    151  AAACAAACCG GCCTCTTGGG CATGATCGGC AAATCCGagc gcgGcgcggc
    201  cacctGCGAA GCcatCGCCG ACAACAAGGC CGTGTACCTC ATGGCAGTCG
    251  GCGGCGCGGC ATACCTCGTG GCAAAAGCCA TCAAATCTTC CAAAGTCTTG
    301  GCGTTCCCCG AATTGGGTAT GGAAGCCGTT TACGAATTTG AAGTCAAAGA
    351  TATGCCCGTA ACCGTCGCCG TGGACAGCAA AGGCGAATCC ATCCACGCCA
    401  CCGCCCCGCG CAAATGGCAG GCGAAAATCG GCATCATCCC CGTCGAGTCT
    451  TGA
```

This corresponds to the amino acid sequence <SEQ ID 200; ORF 048.ng>:

```
g048.pep
      1  MLDKGEELPV DFTNRLIYYV GPVDPVGDEV VGPAGPTTAT RMDKFTRQML
     51  KQTGLLGMIG KSERGAATCE AIADNKAVYL MAVGGAAYLV AKAIKSSKVL
    101  AFPELGMEAV YEFEVKDMPV TVAVDSKGES IHATAPRKWQ AKIGIIPVES
    151  *
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 201>:

```
m048.seq
      1  ATGCTCAACA AAGGCGAAGA ATTGCCCGTC GATTTCACCA ACCGCCTGAT
     51  TTACTACGTC GGCCCCGTCG ATCCGGTCGG CGATGAAGTC GTCGGTCCGG
```

```
101   CAGGTCCGAC CACAGCCACC CGCATGGACA AATTCACCCG CCAAATGCTC
151   GAACAAACCG ACCTCTTGGG CATGATCGGC AAATCCGAGC GCGGCGTGGC
201   CACCTGCGAA GCCATCGCCG ACAACAAAGC CGTGTACCTC ATGGCAGTCG
251   GCGGCGCGGC GTATCTCGTG GCAAAAGCCA TCAAATCTTC CAAAGTCTTG
301   GCGTTCCCCG AATTGGGCAT GGAAGCCATT TACGAATTTG AAGTCAAAGA
351   CATGCCCGTA ACCGTCGCCG TAGATAGCAA AGGCGAATCC ATCCACGCCA
401   CCGCCCCGCG CAAATGGCAG GCGAAAATCG GCATCATCCC CGTCGAATCT
451   TGA
```

This corresponds to the amino acid sequence <SEQ ID 202: ORF 048>:

```
m048.pep
    1   MLNKGEELPV DFTNRLIYYV GPVDPVGDEV VGPAGPTTAT RMDKFTRQML
   51   EQTDLLGMIG KSERGVATCE AIADNKAVYL MAVGGAAYLV AKAIKSSKVL
  101   AFPELGMEAI YEFEVKDMPV TVAVDSKGES IHATAPRKWQ AKIGIIPVES
  151   *
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 203>:

```
a048.seq
    1   ATGCTCGACA AAGGCGAAGA ATTGCCCGTC GATTTCACCA ACCGCCTGAT
   51   TTACTACGTC GGCCCCGTCG ATCCGGTCGG CGACGAAATC GTCGGCCCAG
  101   CAGGTCCGAC CACCGCCACC CGCATGGACA AATTCACCCG CCAAATGCTC
  151   GAACAAACCG ACCTCTTGGG CATGATCGGC AAATCCGAGC GCGGCGCGGC
  201   CACCTGCGAA GCCATCGCCG ACAACAAAGC CGTGTACCTC ATGGCAGTCG
  251   GCGGCGCGGC GTATCTCGTG GCAAAAGCCA TCAAATCTTC CAAAGTCTTG
  301   GCGTTCCCCG AATTGGGCAT GGAAGCCATT TACGAATTTG AAGTCAAAGA
  351   CATGCCCGTA ACCGTCGCCG TAGACAGCAA AGGCGAATCC ATCCACGCCA
  401   CCGCCCCGCC CCAATGGCAG GCGAAAATCG GCATCATCCC CGTCAAATCT
  451   TGA
```

This corresponds to the amino acid sequence <SEQ ID 204; ORF 048.a>:

```
a048.pep
    1   MLDKGEELPV DFTNRLIYYV GPVDPVGDEI VGPAGPTTAT RMDKFTRQML
   51   EQTDLLGMIG KSERGAATCE AIADNKAVYL MAVGGAAYLV AKAIKSSKVL
  101   AFPELGMEAI YEFEVKDMPV TVAVDSKGES IHATAPPQWQ AKIGIIPVKS
  151   *
```

m048/a048 96.0% identity over a 150 aa overlap

```
                     10        20        30        40        50        60
m048.pep     MLNKGEELPVDFTNRLIYYVGPVDPVGDEVVGPAGPTTATRMDKFTRQMLEQTDLLGMIG
             ||:|||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
a048         MLDKGEELPVDFTNRLIYYVGPVDPVGDEIVGPAGPTTATRMDKFTRQMLEQTDLLGMIG
                     10        20        30        40        50        60

                     70        80        90       100       110       120
m048.pep     KSERGVATCEAIADNKAVYLMAVGGAAYLVAKAIKSSKVLAFPELGMEAIYEFEVKDMPV
             |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
a048         KSERGAATCEAIADNKAVYLMAVGGAAYLVAKAIKSSKVLAFPELGMEAIYEFEVKDMPV
                     70        80        90       100       110       120

                    130       140       150
m048.pep     TVAVDSKGESIHATAPRKWQAKIGIIPVESX
             |||||||||||||||| :|||||||||||:||
a048         TVAVDSKGESIHATAPPQWQAKIGIIPVKSX
                    130       140       150
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 048 shows 96.4% identity over a 150 aa overlap with a predicted ORF (ORF 048.ng) from *N. gonorrhoeae:*

```
m048/g048

                       10        20        30        40        50        60
m048.pep    MLNKGEELPVDFTNRLIYYVGPVDPVGDEVVGPAGPTTATRMDKFTRQMLEQTDLLGMIG
            ||:||||||||||||||||||||||||||||||||||||||||||||||:|| |||||||
g048        MLDKGEELPVDFTNRLIYYVGPVDPVGDEVVGPAGPTTATRMDKFTRQMLKQTGLLGMIG
                       10        20        30        40        50        60

                       70        80        90       100       110       120
m048.pep    KSERGVATCEAIADNKAVYLMAVGGAAYLVAKAIKSSKVLAFPELGMEAIYEFEVKDMPV
            |||||:|||||||||||||||||||||||||||||||||||||||||||:||||||||||
g048        KSERGAATCEAIADNKAVYLMAVGGAAYLVAKAIKSSKVLAFPELGMEAVYEFEVKDMPV
                       70        80        90       100       110       120

                      130       140       150
m048.pep    TVAVDSKGESIHATAPRKWQAKIGIIPVESX
            |||||||||||||||||||||||||||||||
g048        TVAVDSKGESIHATAPRKWQAKIGIIPVESX
                      130       140       150
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 205>:

```
g049.seq
      1  ATGCGGGCGC AGGCGTTTGA TCAACCGTTC GGTCAGCTCC TGTTCGGACA
     51  GGCAGAACAC TTCGCGCCGG TTGACGGCTT TCGGGTTCAG GATATTGATT
    101  TGGACGGGCA TCAACGCCTC TTCCGCACCG CCTTCGCCGT TTTCCGCAAC
    151  CCCGTCTGCC GCCGTACCGG ATTCTGCCGC ATCGGCGTTT TCCCCGCCCT
    201  CAATCTGTGC GGTTTCAAAT TCGGCACTGT CTTTTTTGGC ATCGAACCGG
    251  ATTCTCCGCC GCGATTCGAT GTGTTTTTCC GAAAccggca tTTGCAGGGA
    301  AGCCTgcgcg TTGAGCCAGT TTTCCTGAAG GACGATCATC GGGTCGGTTT
    351  CGACTTCCTC GCCGCAATCG GCAACGGCgc tGTTGTGTTC TTCCTGCCAT
    401  TTCTTCAGAT ACGCCTTTAA
```

This corresponds to the amino acid sequence <SEQ ID 206; ORF 049.ng>:

```
g049.pep
      1  MRAQAFDQPF GQLLFGQAEH FAPVDGFRVQ DIDLDGHQRL FRTAFAVFRN
     51  PVCRRTGFCR IGVFPALNLC GFKFGTVFFG IEPDSPPRFD VFFRNRHLQG
    101  SLRVEPVFLK DDHRVGFDFL AAIGNGAVVF FLPFLQIRL*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 207>:

```
m049.seq  (partial)
      1  ATGCGGGCGC AGGCGTTTGA TCAGCCGTTC GGTCAGCTCC TGTTCGGACA
     51  GGCAGAACAC TTCGCGCCGG TTGACGGCTT TCGGGTTCAG GATATTGATT
    101  TGGACGGGCA TCAACGTTTC TTCCGCATCG TTTTCCCCGT TTTCCGAAAC
    151  CGCCGGCTCA TTCGTGCCGG ATTCTGCCTC GTCGGCGTTT TCCCCGCTTT
    201  CAATCTGTCC GGTTTCAAAT TCGACACTGT CTTTTTTGGT ATCAAACCGG
    251  ATTCTCCGCC GCGATTCGAT GTGTTTTTCC GAAACCGACA TTTGCAGGGA
    301  AGCCTGCGCG TTGAGCCAGT TTTCCTGAAG GACGATCATC GGGTCGGTTT
    351  CGACTTCCTC GCCGCAATCG GCAACGGCGG CATTGTGTTC CTCCTGCCAT
    401  TTTTTCAGAT ACGCCTT...
```

This corresponds to the amino acid sequence <SEQ ID 208; ORF 049>:

```
m049.pep  (partial)
      1  MRAQAFDQPF GQLLFGQAEH FAPVDGFRVQ DIDLDGHQRF FRIVFPVFRN
     51  RRLIRAGFCL VGVFPAFNLS GFKFDTVFFG IKPDSPPRFD VFFRNRHLQG
    101  SLRVEPVFLK DDHRVGFDFL AAIGNGGIVF LLPFFQIRL...
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 209>:

```
a049.seq
```

```
  1  ATGCGGGCGC AGGCGTTTGA TCAGCCGTTC GGTCAGCTCC TGTTCGGACA
 51  GGCAGAACAC TTCGCGCCGG TTGACGGCTT TCGGGTTCAG AATATTGATT
101  TGGACGGGCA TCAACGCTTC TTCCGCACCG CCTTCGCCGT TTTCCGCAAC
151  CCCGTCTGCC GCCGTACCCG ATTCTGCCGC ATCGGCGTTT TCCCCGCCTT
201  CAATCTGTCC GGTTTCAAAT TCGGCACTGT CTTTTTTGGC ATCAAACCGG
251  ATTCTCCGCC GCGATTCGAT GTGTTTTTCC GAAACCGACA TTTGCAGGGA
301  AGCCTGCGCG TTGAGCCAGT TTTCCTGAAG GACGATCATC GGGTCGGTTT
351  CGACTTCCTC GCCGCAATCG GCAACGGCGG CATTGTGTTC CTCCTGCCAT
401  TTTTTCAGAT ACGCCTT
```

This corresponds to the amino acid sequence <SEQ ID 210: ORF 049.a>:

```
a049.pep
  1  MRAQAFDQPF GQLLFGQAEH FAPVDGFRVQ NIDLDGHQRF FRTAFAVFRN
 51  PVCRRTRFCR IGVFPAFNLS GFKFGTVFFG IKPDSPPRFD VFFRNRHLQG
101  SLRVEPVFLK DDHRVGFDFL AAIGNGGIVF LLPFFQIRL
```

m049/a049 90.6% identity over a 139 aa overlap

```
                    10        20        30        40        50        60
m049.pep    MRAQAFDQPFGQLLFGQAEHFAPVDGFRVQDIDLDGHQRFFRIVFPVFRNRRLIRAGFCL
            ||||||||||||||||||||||||||||||:||||||||||| :| ||||    |: || 
a049        MRAQAFDQPFGQLLFGQAEHFAPVDGFRVQNIDLDGHQRFFRTAFAVFRNPVCRRTRFCR
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m049.pep    VGVFPAFNLSGFKFDTVFFGIKPDSPPRFDVFFRNRHLQGSLRVEPVFLKDDHRVGFDFL
            :||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
a049        IGVFPAFNLSGFKFGTVFFGIKPDSPPRFDVFFRNRHLQGSLRVEPVFLKDDHRVGFDFL
                    70        80        90       100       110       120

                   130      139
m049.pep    AAIGNGGIVFLLPFFQIRL
            |||||||||||||||||||
a049        AAIGNGGIVFLLPFFQIRL
                   130
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

ORF 049 shows 86.3% identity over a 139 aa overlap with a predicted ORF (ORF 049.ng) from *N. gonorrhoeae:*

```
m049/g049

                    10        20        30        40        50        60
m049.pep    MRAQAFDQPFGQLLFGQAEHFAPVDGFRVQDIDLDGHQRFFRIVFPVFRNRRLIRAGFCL
            |||||||||||||||||||||||||||||||||||||||:|| :| ||||    |:||| 
g049        MRAQAFDQPFGQLLFGQAEHFAPVDGFRVQDIDLDGHQRLFRTAFAVFRNPVCRRTGFCR
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m049.pep    VGVFPAFNLSGFKFDTVFFGIKPDSPPRFDVFFRNRHLQGSLRVEPVFLKDDHRVGFDFL
            :|||||:|| |||| ||||||:||||||||||||||||||||||||||||||||||||||
g049        IGVFPALNLCGFKFGTVFFGIEPDSPPRFDVFFRNRHLQGSLRVEPVFLKDDHRVGFDFL
                    70        80        90       100       110       120

                   130      139
m049.pep    AAIGNGGIVFLLPFFQIRL
            ||||||::||:|||:||||
g049        AAIGNGAVVFFLPFLQIRLX
                   130       140
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 211>:

```
g050.seq
  1   atgggcgCGG GCTGGTGTCC TCCCGGCATC TTGGGCATCG GCATCGGCGg
 51   cacgcccGAA AAAGccgtgt TGATGGcaaA AGAATCCCTG ATGAGCCACA
101   TCGAcatCca aGaATTGCAG GAAAAAGCCG CGTccggggc ggaattgtcc
151   accaccgaAG ccCTGCGCCT cGAACTCTTT GAAAAGGTCA ACGCGCTGGG
201   CATCGGCGCG CAAGGCTTGG GCGGTCTGAC CACCGTGTTG GACGTGAAAA
251   TCCTCGATTA CCCGACCCAT GCCGCCTCCA AACCGATTGC CATGATTCCC
301   AACTGTGCcg ccacCCGcca cgtcgAATTT GAATTGgACG GCTCAGGtcc
351   TGTCGAactc acgccGCcgc gtgtCGAAGA CTGA
```

This corresponds to the amino acid sequence <SEQ ID 212; ORF 050.ng>:

```
g050.pep
  1   MGAGWCPPGI LGIGIGGTPE KAVLMAKESL MSHIDIQELQ EKAASGAELS
 51   TTEALRLELF EKVNALGIGA QGLGGLTTVL DVKILDYPTH AASKPIAMIP
101   NCAATRHVEF ELDGSGPVEL TPPRVED*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 213>:

```
m050.seq
  1   ATGGGCGCGG GCTGGTGTCC TCCCGGCATC TTGGGTATCG GCATCGGCGG
 51   C..agCCgAA AAAGCCGTGC TGATGGCAAA AGAGTCCCTG ATGAGCCACA
101   TCGACATTCA AGAATTGCAG GAAAAGGCCG CGTCCGGCGC GgAATTGTCC
151   ACCACCGAAG CCCTGCGCCT CGAACTCTTT GAAAAAGTCA ACGCGCTGGG
201   CATCGGCGCA CAAGGCTTGG GCGGACTGAC CACCGTGTTG GACGTGAAAA
251   TCCTCGATTA TCCGACCCAC GCCGCCTCCA AACCGATTGC CATGATTCCG
301   AACTGCGCCG CCACCCGCCA CGTCGAATTT GAATTGGACG GCTCAGGCCC
351   TGTCGAACTC ACGCCGCCGC GCGTCGAAGA TGGCCCGATT TGA
```

This corresponds to the amino acid sequence <SEQ ID 214; ORF 050>:

```
m050.pep
  1   MGAGWCPPGI LGIGIGGXAE KAVLMAKESL MSHIDIQELQ EKAASGAELS
 51   TTEALRLELF EKVNALGIGA QGLGGLTTVL DVKILDYPTH AASKPIAMIP
101   NCAATRHVEF ELDGSGPVEL TPPRVEDGPI *
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 215>:

```
a050.seq
  1   ATGGGCGCGG GCTGGTGTCC TCCCGGCATC TTGGGCATCG GCATCGGCGG
 51   TACGCCCGAA AAAGCCGTGT TGATGGCGAA AGAATCCCTG ATGAGCCACA
101   TCGACATCCA AGAATTGCAG GAAAAAGCCG CGTCCGGCGC GGAATTGTCC
151   ACCACCGAAG CCCTGCGCCT CGAACTCTTT GAAAAAGTCA ACGCGCTAGG
201   CATCGGCGCG CAAGGCTTGG GCGGTCTGAC CACCGTGTTG GACGTGAAAA
251   TCCTCGATTA CCCGACCCAC GCCGCCTCCA AACCGATTGC CATGATTCCG
301   AACTGCGCCG CCACCCGCCA CGTCGAATTT GAATTGGACG GCTCAGGCCC
351   TGTCGAACTC ACGCCGCCGC GCGTCGAAGA CTGGCCC
```

This corresponds to the amino acid sequence <SEQ ID 216; ORF 050.a>:

```
a050.pep
  1   MGAGWCPPGI LGIGIGGTPE KAVLMAKESL MSHIDIQELQ EKAASGAELS
 51   TTEALRLELF EKVNALGIGA QGLGGLTTVL DVKILDYPTH AASKPIAMIP
101   NCAATRHVEF ELDGSGPVEL TPPRVEDWP
```

m050/a050 97.7% identity over a 129 aa overlap

```
                       10        20        30        40        50        60
m050.pep    MGAGWCPPGILGIGIGGXAEKAVLMAKESLMSHIDIQELQEKAASGAELSTTEALRLELF
            |||||||||||||||||: |||||||||||||||||||||||||||||||||||||||||
a050        MGAGWCPPGILGIGIGGTPEKAVLMAKESLMSHIDIQELQEKAASGAELSTTEALRLELF
```

```
                  10        20        30        40        50        60

                  70        80        90       100       110       120
m050.pep    EKVNALGIGAQGLGGLTTVLDVKILDYPTHAASKPIAMIPNCAATRHVEFELDGSGPVEL
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a050        EKVNALGIGAQGLGGLTTVLDVKILDYPTHAASKPIAMIPNCAATRHVEFELDGSGPVEL
                  70        80        90       100       110       120

                 130
m050.pep    TPPRVEDGPIX
            |||||||  |
a050        TPPRVEDWP
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 050 shows 98.4% identity over a 127 aa overlap with a predicted ORF (ORF 050.ng) from *N. gonorrhoeae*:

m050/g050

```
                  10        20        30        40        50        60
m050.pep    MGAGWCPPGILGIGIGGXAEKAVLMAKESLMSHIDIQELQEKAASGAELSTTEALRLELF
            ||||||||||||||||: ||||||||||||||||||||||||||||||||||||||||||
g050        MGAGWCPPGILGIGIGGTPEKAVLMAKESLMSHIDIQELQEKAASGAELSTTEALRLELF
                  10        20        30        40        50        60

                  70        80        90       100       110       120
m050.pep    EKVNALGIGAQGLGGLTTVLDVKILDYPTHAASKPIAMIPNCAATRHVEFELDGSGPVEL
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g050        EKVNALGIGAQGLGGLTTVLDVKILDYPTHAASKPIAMIPNCAATRHVEFELDGSGPVEL
                  70        80        90       100       110       120

                 130
m050.pep    TPPRVEDGPIX
            |||||||
g050        TPPRVEDX
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 217>:

g050-1.seq

```
   1  ATGACCGTTA TCAAGCAAGA AGACTTTATT CAAAGTATCT GCGATGCCTT
  51  CCAATTCATC AGCTACTACC ATCCAAAAGA CTACATCGAC GCGCTTTATA
 101  AGGCGTGGCA GAAGGAAGAA AATCCCGCCG CCAAAGACGC GATGACGCAG
 151  ATTTTGGTCA ACAGCCGTAT GTGTGCCGAA AACAACCGCC CCATCTGCCA
 201  AGACACAGGT ATCGCAACCG TCTTCCTCAA AGTCGGTATG GATGTGCAAT
 251  GGGATGCGGA CATGAGCGTG GAAAAGATGG TTAACGAAGG CGTACGCCGC
 301  GCCTACACTT GGGAAGGCAA CACCCTGCGC GCTTCCGTCC TCGCCGATCC
 351  GGCCGGCAAA CGCCAAAACA CCAAAGACAA CACCCCCGCC GTCATCCACA
 401  TGAGCATCGT GCCGGGCGGT AAAGTCGAAG TAACCTGCGC GGCAAAAGGC
 451  GGCGGCTCTG AAAACAAATC CAAACTCGCT ATGCTCAACC CTTCCGACAA
 501  CATCGTCGAT TGGGTATTGA AAACCATCCC GACGATGGGC GCGGGCTGGT
 551  GTCCTCCCGG CATCTTGGGC ATCGGCATCG GCGGCAcgcC CGAAAAAGCC
 601  GTGTTGATGG cgaAAGAATC CCTGATGAGC CACATCGACA TCCAAGAATT
 651  GCAGGAAAAA GCCGCGTCCG GCGCGGAATT GTCCACCACC GAAGCCCTGC
 701  GCCTCGAACT CTTTGAAAAG GTCAACGCGC TGGGCATCGG CGCGCAAGGC
 751  TTGGGCGGTC TGACCACCGT GTTGGACGTG AAAATCCTCG ATTACCCGAC
 801  CCATGCCGCC TCCAAACCGA TTGCCATGAT TCCCAACTGT GCCGCCACCC
 851  GCCACGTCGA ATTTGAATTG GACGGCTCAG GTCCTGTCGA ACTCACGCCG
 901  CCGCGCGTCG AAGACTGACC CGATCTGACT TACAGCCCCG ACAACGGCAA
 951  ACGCGTCGAT GTCGATAAGC TGACCAAAGA AGAAGTGGCA AGCTGGAAAA
1001  CCGGCGACGT ATTGCTGTTG AACGGCAAAA TCCTCACCCG CCGCGATGCC
1051  GCGCACAAAC GCCTCGTCAA TATGCTCGAC AAAGGCGAGG AGTTGCCCGT
1101  CGATTTCACC AACCGCCTGA TTTACTACGT CGGCCCCGTC GATCCGGTCG
1151  GCGATGAAGT CGTCGGTCCC GCAGGTCCGA CCACAGCCAC CCGCATGGAC
1201  AAATTTACCC GCCAAATGCT CAAACAAACC GGCCTCTTGG GCATGATCGG
```

```
1251  CAAATCCGAG CGCGGCGCGG CCACCTGCGA AGCCATCGCC GACAACAAGG
1301  CCGTGTACCT CATGGCAGTC GGCGGCGCGG CATACCTCGT GGCAAAAGCC
1351  ATCAAATCTT CCAAAGTCTT GGCGTTCCCC GAATTGGGTA TGGAAGCCGT
1401  TTACGAATTT GAAGTCAAAG ATATGCCCGT AACCGTCGCC GTGGACAGCA
1451  AAGGCGAATC CATCCACGCC ACCGCCCCGC GCAAATGGCA GGCGAAAATC
1501  GGCATCATCC CCGTCGAGTC TTGA
```

This corresponds to the amino acid sequence <SEQ ID 218; ORF 050-1.ng>:

g050-1.pep

```
  1  MTVIKQEDFI QSICDAFQFI SYYHPKDYID ALYKAWQKEE NPAAKDAMTQ
 51  ILVNSRMCAE NNRPICQDTG IATVFLKVGM DVQWDADMSV EKMVNEGVRR
101  AYTWEGNTLR ASVLADPAGK RQNTKDNTPA VIHMSIVPGG KVEVTCAAKG
151  GGSENKSKLA MLNPSDNIVD WVLKTIPTMG AGWCPPGILG IGIGGTPEKA
201  VLMAKESLMS HIDIQELQEK AASGAELSTT EALRLELFEK VNALGIGAQG
251  LGGLTTVLDV KILDYPTHAA SKPIAMIPNC AATRHVEFEL DGSGPVELTP
301  PRVED*PDLT YSPDNGKRVD VDKLTKEEVA SWKTGDVLLL NGKILTGRDA
351  AHKRLVNMLD KGEELPVDFT NRLIYYVGPV DPVGDEVVGP AGPTTATRMD
401  KFTRQMLKQT GLLGMIGKSE RGAATCEAIA DNKAVYLMAV GGAAYLVAKA
451  IKSSKVLAFP ELGMEAVYEF EVKDMPVTVA VDSKGESIHA TAPRKWQAKI
501  GIIPVES*
```

g050-1/p14407

```
 sp|P14407|FUMB_ECOLI FUMARATE HYDRATASE CLASS I, ANAEROBIC (FUMARASE)
>gi|280063|pir||B44511 fumarate hydratase (EC 4.2.1.2) fumB, iron-dependent - Escherichia coli
>gi|146048 (M27058) anaerobic class I fumarase (EC 4.2.1.2) [Escherichia coli] Length = 548
 Score =  172 bits (432), Expect = 4e-42
 Identities = 138/488 (28%), Positives = 216/488 (43%), Gaps = 22/488 (4%)

Query: 11   QSICDAFQFISYYHPKDYIDALYKAWQKEENPAAKDAMTQILVNSRMCAENNRPICQDTG 70
            Q+  DA    +  H K      L+       E +     K      Q L NS + A+    P CQDTG
Sbjct: 53   QAFHDASFMLRPAHQKQVAAILHDPEASEND---KYVALQFLRNSEIAAKGVLPTCQDTG 109

Query: 71   IATVFLKVGMDVQWDADMSVEKMVNEGVRRAYTWEGNTLRASVLADPAGKRQNTKDNTPA 130
             A +  K G  V W           E+ +++GV   Y  E N   +   A      K  NT  N PA
Sbjct: 110  TAIIVGKKGQRV-WTGGGD-EETLSKGVYNTYI-EDNLRYSQNAALDMYKEVNTGTNLPA 166

Query: 131  VIHMSIVPGGKVEVTCAAKGGGSENKSKL-----AMLNPSDNIVDWVLKTIPTMGAGWCP 185
             I +  V G + +  C AKGGGS NK+ L       A+L P   + ++++ + T+G    CP
Sbjct: 167  QIDLYAVDGDEYKFLCVAKGGGSANKTYLYQETKALLTPG-KLKNFLVEKMRTLGTAACP 225

Query: 186  PXXXXXXXXXTPEKAVLMAKESLMSHIDIQELQEKAASGAELSTTEALRLELFEKVNXXX 245
            P           T  + L    +    +H       EL  +         +         L  EL E+
Sbjct: 226  PYHIAFVIGGTSAETNLKTVKLASAHY-YDELPTEGNEHGQAFRDVQLEQELLEEAQKLG 284

Query: 246  XXXXXXXXXTTVLDVKILDYPTHAASKPIAMIPNCAATRHVEFELDGSG----PVELTPP 301
                           D++++  P H AS P+ M  +C+A R+++ +++   G      +E  P
Sbjct: 285  LGAQFGGKYFAH-DIRVIRLPRHGASCPVGMGVSCSADRNIKAKINREGIWIEKLEHNPG 343

uery: 302   RVEDXPDLTYSPDNGKRVDVDKLTKE---EVASWKTGDVLLLNGKILTGRDAAHKRLVNM 358
          +                +VD+++   KE     +++ +       L  L  G I+ GRD AH +L  +
Sbjct: 344  QYIPQELRQAGEGEAVKVDLNRPMKEILAQLSQYPVSTRLSLTGTIIVGRDIAHAKLKEL 403

Query: 359  LDKGEELPVDFTNRLIYYXXXXXXXXXXXXXXXXXXTTATRMDKFTRQMLKQTGLLGMIGK 418
            +D G+ELP     +   IYY                          TTA RMD +    +     G + M+ K
Sbjct: 404  IDAGKELPQYIKDHPIYYAGPAKTPAGYPSGSLGPTTAGRMDSYVDLLQSHGGSMIMLAK 463

Query: 419  SERGAATCEAIADNKAVYLMAVGG-AAYLVAKAIKSSKVLAFPELGMEAVYEFEVKDMPV 477
             R       +A    +   YL ++GG AA L   ++IK  + +A+PELGMEA+++ EV+D P
Sbjct: 464  GNRSQQVTDACHKHGGFYLGSIGGPAAVLAQQSIKHLECVAYPELGMEAIWKIEVEDFPA 523

Query: 478  TVAVDSKG 485
             + VD KG
Sbjct: 524  FILVDDKG 531
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 219>:

m050-1.seq

```
  1  ATGACCGTCA TCAAACAGGA AGACTTTATC CAAAGCATTT GCGATGCCTT
 51  CCAATTCATC AGCTACTATC ATCCCAAAGA CTACATCGAC GCGCTTTATA
101  AGGCGTGGCA GAAGGAAGAA AATCCTGCCG CCAAAGACGC GATGACGCAG
151  ATTTTGGTCA ACAGCCGTAT GTGTGCGGAA AACAACCGCC CCATCTGCCA
201  AGACACAGGT ATCGCAACCG TCTTCCTCAA AGTCGGTATG AACGTCCAAT
251  GGGATGCGGA CATGAGCGTG GAAGAGATGG TTAACGAAGG CGTACGCCGC
```

```
 301  GCCTACACTT GGGAAGGCAA TACGCTGCGC GCTTCCGTCC TCGCCGATCC
 351  GGCCGGCAAA CGCCAAAACA CCAAAGACAA CACCCCCGCC GTCATCCATA
 401  TGAGCATCGT GCCGGGCGGT AAAGTCGAAG TAACCTGCGC GGCAAAAGGC
 451  GGCGGCTCTG AAAACAAATC CAAACTCGCC ATGCTCAATC CTTCCGACAA
 501  CATCGTCGAT TGGGTATTGA AAACCATCCC GACCATGGGC GCGGGCTGGT
 551  GTCCTCCCGG CATCTTGGGT ATCGGCATCG GCGGCACGCC CGAAAAAGCC
 601  GTGCTGATGG CAAAAGAGTC CCTGATGAGC CACATCGACA TTCAAGAATT
 651  GCAGGAAAAG GCCGCGTCCG GCGCGGAATT GTCCACCACC GAAGCCCTGC
 701  GCCTCGAACT CTTTGAAAAA GTCAACGCGC TGGGCATCGG CGCACAAGGC
 751  TTGGGCGGAC TGACCACCGT GTTGGACGTG AAAATCCTCG ATTATCCGAC
 801  CCACGCCGCC TCCAAACCGA TTGCCATGAT TCCGAACTGC GCCGCCACCC
 851  GCCACGTCGA ATTTGAATTG GACGGCTCAG GCCCTGTCGA ACTCACGCCG
 901  CCGCGCGTCG AAGACTGGCC CGATTTGACT TACAGCCCCG ACAACGGCAA
 951  ACGCGTCGAT GTCGACAAGC TGACCAAAGA AGAAGTGGCA AGCTGGAAAA
1001  CCGGCGACGT ATTGCTGTTG AACGGCAAAA TCCTCACCGG CCGCGATGCC
1051  GCACACAAAC GCCTCGTCGA TATGCTCAAC AAAGGCGAAG AATTGCCCGT
1101  CGATTTCACC AACCGCCTGA TTTACTACGT CGGCCCCGTC GATCCGGTCG
1151  GCGATGAAGT CGTCGGTCCG GCAGGTCCGA CCACAGCCAC CCGCATGGAC
1201  AAATTCACCC GCCAAATGCT CGAACAAACC GACCTCTTGG GCATGATCGG
1251  CAAATCCGAG CGCGGCGTGG CCACCTGCGA AGCCATCGCC GACAACAAAG
1301  CCGTGTACCT CATGGCAGTC GGCGGCGCGG CGTATCTCGT GGCAAAAGCC
1351  ATCAAATCTT CCAAAGTCTT GGCGTTCCCC GAATTGGGCA TGGAAGCCAT
1401  TTACGAATTT GAAGTCAAAG ACATGCCCGT AACCGTCGCC GTAGATAGCA
1451  AAGGCGAATC CATCCACGCC ACCGCCCCGC GCAAATGGCA GGCGAAAATC
1501  GGCATCATCC CCGTCGAATC TTGA
```

This corresponds to the amino acid sequence <SEQ ID 220; ORF 050-1>:

m050-1.pep

```
  1  MTVIKQEDFI QSICDAFQFI SYYHPKDYID ALYKAWQKEE NPAAKDAMTQ
 51  ILVNSRMCAE NNRPICQDTG IATVFLKVGM NVQWDADMSV EEMVNEGVRR
101  AYTWEGNTLR ASVLADPAGK RQNTKDNTPA VIHMSIVPGG KVEVTCAAKG
151  GGSENKSKLA MLNPSDNIVD WVLKTIPTMG AGWCPPGILG IGIGGTPEKA
201  VLMAKESLMS HIDIQELQEK AASGAELSTT EALRLELFEK VNALGIGAQG
251  LGGLTTVLDV KILDYPTHAA SKPIAMIPNC AATRHVEFEL DGSGPVELTP
301  PRVEDWPDLT YSPDNGKRVD VDKLTKEEVA SWKTGDVLLL NGKILTGRDA
351  AHKRLVDMLN KGEELPVDFT NRLIYYVGPV DPVGDEVVGP AGPTTATRMD
401  KFTAQMLEQT DLLGMIGKSE RGVATCEAIA DNKAVYLMAV GGAAYLVAKA
451  IKSSKVLAFP ELGMEAIYEF EVKDMPVTVA VDSKGESIHA TAPRKWQAKI
501  GIIPVES*
```

m050-1/g050-1 98.2% identity in 507 aa overlap

```
                     10        20        30        40        50        60
m050-1.pep   MTVIKQEDFIQSICDAFQFISYYHPKDYIDALYKAWQKEENPAAKDAMTQILVNSRMCAE
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g050-1       MTVIKQEDFIQSICDAFQFISYYHPKDYIDALYKAWQKEENPAAKDAMTQILVNSRMCAE
                     10        20        30        40        50        60

                     70        80        90       100       110       120
m050-1.pep   NNRPICQDTGIATVFLKVGMNVQWDADMSVEEMVNEGVRRAYTWEGNTLRASVLADPAGK
             ||||||||||||||||||||:||||||||||:||||||||||||||||||||||||||||
g050-1       NNRPICQDTGIATVFLKVGMDVQWDADMSVEKMVNEGVRRAYTWEGNTLRASVLADPAGK
                     70        80        90       100       110       120

                    130       140       150       160       170       180
m050-1.pep   RQNTKDNTPAVIHMSIVPGGKVEVTCAAKGGGSENKSKLAMLNPSDNIVDWVLKTIPTMG
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g050-1       RQNTKDNTPAVIHMSIVPGGKVEVTCAAKGGGSENKSKLAMLNPSDNIVDWVLKTIPTMG
                    130       140       150       160       170       180

                    190       200       210       220       230       240
m050-1.pep   AGWCPPGILGIGIGGTPEKAVLMAKESLMSHIDIQELQEKAASGAELSTTEALRLELFEK
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g050-1       AGWCPPGILGIGIGGTPEKAVLMAKESLMSHIDIQELQEKAASGAELSTTEALRLELFEK
                    190       200       210       220       230       240

                    250       260       270       280       290       300
m050-1.pep   VNALGIGAQGLGGLTTVLDVKILDYPTHAASKPIAMIPNCAATRHVEFELDGSGPVELTP
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g050-1       VNALGIGAQGLGGLTTVLDVKILDYPTHAASKPIAMIPNCAATRHVEFELDGSGPVELTP
                    250       260       270       280       290       300
```

```
                    310       320       330       340       350       360
m050-1.pep   PRVEDWPDLTYSPDNGKRVDVDKLTKEEVASWKTGDVLLLNGKILTGRDAAHKRLVDMLN
             ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||:||:
g050-1       PRVEDXPDLTYSPDNGKRVDVDKLTKEEVASWKTGDVLLLNGKILTGRDAAHKRLVNMLD
                    310       320       330       340       350       360

                    370       380       390       400       410       420
m050-1.pep   KGEELPVDFTNRLIYYVGPVDPVGDEVVGPAGPTTATRMDKFTRQMLEQTDLLGMIGKSE
             |||||||||||||||||||||||||||||||||||||||||||||||:|| |||||||||
g050-1       KGEELPVDFTNRLIYYVGPVDPVGDEVVGPAGPTTATRMDKFTRQMLKQTGLLGMIGKSE
                    370       380       390       400       410       420

                    430       440       450       460       470       480
m050-1.pep   RGVATCEAIADNKAVYLMAVGGAAYLVAKAIKSSKVLAFPELGMEAIYEFEVKDMPVTVA
             ||:|||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
g050-1       RGAATCEAIADNKAVYLMAVGGAAYLVAKAIKSSKVLAFPELGMEAVYEFEVKDMPVTVA
                    430       440       450       460       470       480

                    490       500
m050-1.pep   VDSKGESIHATAPRKWQAKIGIIPVESX
             ||||||||||||||||||||||||||||
g050-1       VDSKGESIHATAPRKWQAKIGIIPVESX
                    490       500
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 221>:

a050-1.seq

```
   1  ATGACCGTCA TCAAACAGGA AGACTTTATC CAAAGCATTT GCGATGCCTT
  51  CCAATTCATC AGCTACTACC ATCCCAAAGA CTACATCGAC GCGCTTTATA
 101  AGGCGTGGCA GAAGGAAGAA AACCCCGCCG CCAAAGACGC GATGACGCAG
 151  ATTTTGGTCA ACAGCCGCAT GTGTGCCGAA AACAACCGCC CCATCTGCCA
 201  AGATACCGCT ATCGCGACCG TGTTTTTGAA AGTCGGTATG GATGTGCAAT
 251  GGGATGCAGA CATGAGCGTC GAAGAGATGG TTAACGAAGG CGTGCGCCGC
 301  GCCTACACTT GGGAAGGCAA TACGCTGCGC GCTTCCGTTC TCGCCGACCC
 351  CGCCGGCAAA CGCCAAAATA CCAAAGACAA CACGCCCGCC GTCATCCATA
 401  TGAGCATCGT GCCGGGCGAC AAAGTCGAAG TAACCTGCGC GGCAAAAGGC
 451  GGCGGTTCTG AAAACAAATC CAAACTCGCC ATGCTCAACC CTTCCGACAA
 501  CATCGTCGAT TGGGTATTGA AAACCATTCC GACCATGGGC GCGGGCTGGT
 551  GTCCTCCCGG CATCTTGGGC ATCGGCATCG GCGGTACGCC CGAAAAAGCC
 601  GTGTTGATGG CGAAAGAATC CCTGATGAGC CACATCGACA TCCAAGAATT
 651  GCAGGAAAAA GCCGCGTCCG GCGCGGAATT GTCCACCACC GAAGCCCTGC
 701  GCCTCGAACT CTTTGAAAAA GTCAACGCGC TAGGCATCGG CGCGCAAGGC
 751  TTGGGCGGTC TGACCACCGT GTTGGACGTG AAAATCCTCG ATTACCCGAC
 801  CCACGCCGCC TCCAAACCGA TTGCCATGAT TCCGAACTGC GCCGCCACCC
 851  GCCACGTCGA ATTTGAATTG GACGGCTCAG GCCCTGTCGA ACTCACGCCG
 901  CCGCGCGTCG AAGACTGGCC CGATTTGACT TACAGCCCCG ACAACGGCAA
 951  ACGCGTCGAT GTCGACAAGC TGACCAAAGA AGAAGTGGCA AGCTGGAAAA
1001  CCGGCGACGT ATTGCTGTTG AACGGCAAAA TCCTCACCGG CCGCGATGCC
1051  GCACACAAAC GCCTCGTCGA TATGCTCGAC AAAGGCGAAG AATTGCCCGT
1101  CGATTTCACC AACCGCCTGA TTTACTACGT CGGCCCCGTC GATCCGGTCG
1151  GCGACGAAAT CGTCGGCCCA GCAGGTCCGA CCACCGCCAC CCGCATGGAC
1201  AAATTCACCC GCCAAATGCT CGAACAAACC GACCTCTTGG GCATGATCGG
1251  CAAATCCGAG CGCGGCGCGG CCACCTGCGA AGCCATCGCC GACAACAAAG
1301  CCGTGTACCT CATGGCAGTC GGCGGCGCGG CGTATCTCGT GGCAAAAGCC
1351  ATCAAATCTT CCAAAGTCTT GGCGTTCCCC GAATTGGGCA TGGAAGCCAT
1401  TTACGAATTT GAAGTCAAAG ACATGCCCGT AACCGTCGCC GTAGACAGCA
1451  AAGGCGAATC CATCCACGCC ACCGCCCCGC CCCAATGGCA GGCGAAAATC
1501  GGCATCATCC CCGTCAAATC TTGA
```

This corresponds to the amino acid sequence <SEQ ID 222; ORF 050-1.a>:

a050-1.pep

```
   1  MTVIKQEDFI QSICDAFQFI SYYHPKDYID ALYKAWQKEE NPAAKDAMTQ
  51  ILVNSRMCAE NNRPICQDTG IATVFLKVGM DVQWDADMSV EEMVNEGVRR
 101  AYTWEGNTLR ASVLADPAGK RQNTKDNTPA VIHMSIVPGD KVEVTCAAKG
 151  GGSENKSKLA MLNPSDNIVD WVLKTIPTMG AGWCPPGILG IGIGGTPEKA
 201  VLMAKESLMS HIDIQELQEK AASGAELSTT EALRLELFEK VNALGIGAQG
 251  LGGLTTVLDV KILDYPTHAA SKPIAMIPNC AATRHVEFEL DGSGPVELTP
 301  PRVEDWPDLT YSPDNGKRVD VDKLTKEEVA SWKTGDVLLL NGKILTGRDA
 351  AHKRLVDMLD KGEELPVDFT NRLIYYVGPV DPVGDEIVGP AGPTTATRMD
 401  KFTRQMLEQT DLLGMIGKSE RGAATCEAIA DNKAVYLMAV GGAAYLVAKA
 451  IKSSKVLAFP ELGMEAIYEF EVKDMPVTVA VDSKGESIHA TAPPQWQAKI
 501  GIIPVKS*
```

a050-1/m050-1 98.4% identity in 507 aa overlap

```
                   10        20        30        40        50        60
a050-1.pep  MTVIKQEDFIQSICDAFQFISYYHPKDYIDALYKAWQKEENPAAKDAMTQILVNSRMCAE
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m050-1      MTVIKQEDFIQSICDAFQFISYYHPKDYIDALYKAWQKEENPAAKDAMTQILVNSRMCAE
                   10        20        30        40        50        60

                   70        80        90       100       110       120
a050-1.pep  NNRPICQDTGIATVFLKVGMDVQWDADMSVEEMVNEGVRRAYTWEGNTLRASVLADPAGK
            ||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
m050-1      NNRPICQDTGIATVFLKVGMNVQWDADMSVEEMVNEGVRRAYTWEGNTLRASVLADPAGK
                   70        80        90       100       110       120

                  130       140       150       160       170       180
a050-1.pep  RQNTKDNTPAVIHMSIVPGDKVEVTCAAKGGGSENKSKLAMLNPSDNIVDWVLKTIPTMG
            ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
m050-1      RQNTKDNTPAVIHMSIVPGGKVEVTCAAKGGGSENKSKLAMLNPSDNIVDWVLKTIPTMG
                  130       140       150       160       170       180

                  190       200       210       220       230       240
a050-1.pep  AGWCPPGILGIGIGGTPEKAVLMAKESLMSHIDIQELQEKAASGAELSTTEALRLELFEK
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m050-1      AGWCPPGILGIGIGGTPEKAVLMAKESLMSHIDIQELQEKAASGAELSTTEALRLELFEK
                  190       200       210       220       230       240

                  250       260       270       280       290       300
a050-1.pep  VNALGIGAQGLGGLTTVLDVKILDYPTHAASKPIAMIPNCAATRHVEFELDGSGPVELTP
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m050-1      VNALGIGAQGLGGLTTVLDVKILDYPTHAASKPIAMIPNCAATRHVEFELDGSGPVELTP
                  250       260       270       280       290       300

                  310       320       330       340       350       360
a050-1.pep  PRVEDWPDLTYSPDNGKRVDVDKLTKEEVASWKTGDVLLLNGKILTGRDAAHKRLVDMLD
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
m050-1      PRVEDWPDLTYSPDNGKRVDVDKLTKEEVASWKTGDVLLLNGKILTGRDAAHKRLVDMLN
                  310       320       330       340       350       360

                  370       380       390       400       410       420
a050-1.pep  KGEELPVDFTNRLIYYVGPVDPVGDEIVGPAGPTTATRMDKFTRQMLEQTDLLGMIGKSE
            ||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||
m050-1      KGEELPVDFTNRLIYYVGPVDPVGDEVVGPAGPTTATRMDKFTRQMLEQTDLLGMIGKSE
                  370       380       390       400       410       420

                  430       440       450       460       470       480
a050-1.pep  RGAATCEAIADNKAVYLMAVGGAAYLVAKAIKSSKVLAFPELGMEAIYEFEVKDMPVTVA
            ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m050-1      RGVATCEAIADNKAVYLMAVGGAAYLVAKAIKSSKVLAFPELGMEAIYEFEVKDMPVTVA
                  430       440       450       460       470       480

                  490       500
a050-1.pep  VDSKGESIHATAPPQWQAKIGIIPVKSX
            ||||||||||||| :||||||||||:||
m050-1      VDSKGESIHATAPRKWQAKIGIIPVESX
                  490       500
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 223>:

```
g052.seq
      1  ATGGCTTTGG TGGCGGAGGA AACGGAAATA TCCGCGCCGT GTTTCAAAGG
     51  CTGCGAGCCG ACGGGCGACA GCAGGCTGTT GTCCACCACC AAGAGCGCGC
    101  CGATGCCGTG CGCCAATTCC GCCAAGGCTT CCAAGTCGGC CACTTCGCCC
    151  AAGGGGTTGG ACGGCGTTTC CAAAAACAGC AGTTTGGTGT TGGCTTTGAC
    201  GGCGGCTTTC CATTCATTTA TATCAGTCGG CGACACGCGG CTCACTCCGA
    251  TGCCGAATTT GGTAACGATG TTATTGATAA AGCCGACGGT CGTGCCGAAC
    301  AGGCTGCGGC TGGAAACCAC ATGGTCGCCC GCCTGCAGGA AGGTGAAAAA
    351  CGCCGCCTGA
```

This corresponds to the amino acid sequence <SEQ ID 224; ORF 052.ng>:

g052.pep

```
  1  MALVAEETEI SAPCFKGCEP TGDSRLLSTT KSAPMPCANS AKASKSATSP
 51  KGLDGVSKNS SLVLALTAAF HSFISVGDTR LTPMPNLVTM LLIKPTVVPN
101  RLRLETTWSP ACRKVKNAA*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 225>:

```
m052.seq
  1  ATGGCTTTGG TGGCGGAGGA AACGGAAATA TCCGCGCCGT GTTTCAAAGG
 51  CTGCGAGCCG ACGGGCGACA GCAGGCTGTT GTCCACCACC AAGAGCGCGC
101  CGATGCCGTG CGCCAATTCC GCCAAGGCTT CCAAGTCGGC CACTTCGCCC
151  AAGGGGTTGG ACGGCGTTTC CAAAAACAGC AGTTTGGTGT TGGCTTTGAC
201  GGCGGCTTTC CATTCATTTA TATCAGTCGG CGACACGCGG CTCACTCCGA
251  TGCCGAATTT GGTAACGATG TTATTGATAA AGCCGACGGT CGTGCCGAAC
301  AGGCTGCGGC TGGAAACCAC ATGGTCGCCC GCCTGCAGGA AGGTGAAAAA
351  CGCCGCCTGA
```

This corresponds to the amino acid sequence <SEQ ID 226; ORF 052>:

```
m052.pep
  1  MALVAEETEI SAPCFKGCEP TGDSRLLSTT KSAPMPCANS AKASKSATSP
 51  KGLDGVSKNS SLVLALTAAF HSFISVGDTR LTPMPNLVTM LLIKPTVVPN
101  RLRLETTWSP ACRKVKNAA*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 227>:

```
a052.seq
  1  ATGGCTTTGG TCGCGGAGGA AACGGAAATA TCCGCGCCGT GTTTCAAAGG
 51  CTGAGAGCCG ACAGGCGACA GCAGGCTGTT GTCCACCACC AAGAGCGCGC
101  CGATGCCGTG CGCCAATTCC GCCAAGGCTT CCAAGTCGGC CACTTCTCCC
151  AAGGGATTGG ACGGCGTTTC CAAAAACAGC AGTTTGGTGT TGGCTTTGAC
201  GGCGGCTTTC CATTCGTTTA TATCAGTCGG CGACACGTGA CTCACTTCGA
251  TGCCGAATTT GGTAACGATG TTATTGATAA AGCCGACGGT CGTGCCGAAC
301  AGGCTGCGGC TGGAAATCAC ATGGTCGCCC GCCTGCAAAA AGGTGAAAAA
351  CGCCGCCTGA
```

This corresponds to the amino acid sequence <SEQ ID 228; ORF 052.a>:

```
a052.pep
  1  MALVAEETEI SAPCFKG*EP TGDSRLLSTT KSAPMPCANS AKASKSATSP
 51  KGLDGVSKNS SLVLALTAAF HSFISVGDT* LTSMPNLVTM LLIKPTVVPN
101  RLRLEITWSP ACKKVKNAA*
```

m052/a052 95.8% identity over a 119 aa overlap

```
                    10        20        30        40        50        60
m052.pep    MALVAEETEISAPCFKGCEPTGDSRLLSTTKSAPMPCANSAKASKSATSPKGLDGVSKNS
            ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
a052        MALVAEETEISAPCFKGXEPTGDSRLLSTTKSAPMPCANSAKASKSATSPKGLDGVSKNS
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m052.pep    SLVLALTAAFHSFISVGDTRLTPMPNLVTMLLIKPTVVPNRLRLETTWSPACRKVKNAAX
            ||||||||||||||||||| || |||||||||||||||||||||| ||||||:|||||||
a052        SLVLALTAAFHSFISVGDTXLTSMPNLVTMLLIKPTVVPNRLRLEITWSPACKKVKNAAX
                    70        80        90       100       110       120
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 052 shows 95.8% identity over a 119 aa overlap with a predicted ORF (ORF 052.ng) from *N. gonorrhoeae:*

```
m052/g052

                    10        20        30        40        50        60
m052.pep    MALVAEETEISAPCFKGCEPTGDSRLLSTTKSAPMPCANSAKASKSATSPKGLDGVSKNS
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||
```

```
g052        MALVAEETEISAPCFKGCEPTGDSRLLSTTKSAPMPCANSAKASKSATSPKGLDGVSKNS
                      10        20        30        40        50        60

                      70        80        90       100       110       120
m052.pep    SLVLALTAAFHSFISVGDTWLTSMPNLATMLLIKPTVVPNRLRLEITWSPACKKVKNAAX
            |||||||||||||||||| || ||||:||||||||||||||||,|| ||||||:|||||||
g052        SLVLALTAAFHSFISVGDTRLTPMPNLVTMLLIKPTVVPNRLRLETTWSPACRKVKNAAX
                      70        80        90       100       110       120
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 229>:

```
g073.seq
      1  ATGTGTATGC CATACGCAAT AAGGGTTTCA GACGGCATCT GCCGCATTTT
     51  TCCGCCGATG CCGTCTGAAA CACGCAATCA GCGCGCGAGT GCCTGTTTCA
    101  AATCGTCAAT CAAATCGCCA ACATATTCCA AACCGACCGA CAGGCGCACC
    151  AGTCCGGGGC GGatacCGGC GGCGAGTTTT TCTTCGGGCT GCATCCTGCC
    201  GTGCGTGGTT GTCCACGGAT TGGTGATGGT CGAGCGCACG TCGCCGAGGT
    251  TGGCGGTACG GGAAAAGAGT TCCACGACTT TCCACGCGGC TGCTTGGTCG
    301  GCGACTTCAA AACCGATGAC GATGCCGCCG CCGTTTTGCT GTTTGCGGAT
    351  AAGCTCCGCC TGCGGATGGT CGGGCAATCC GGTGTAG
```

This corresponds to the amino acid sequence <SEQ ID 230; ORF 073.ng>:

```
g073.pep
      1  MCMPYAIRVS DGICRIFPPM PSETRNQRAS ACFKSSIKSP TYSKPTDRRT
     51  SPGRIPAASF SSGCILPCVV VHGLVMVERT SPRLAVREKS STTFHAAAWS
    101  ATSKPMTMPP PFCCLRISSA CGWSGNPV*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 231>:

```
m073.seq
      1  ATGTGTATGC CATATAAGAT AAGGGTTTCA GACGGCATCT GCTGTCCAAT
     51  GCCGTCTGAA ACACGCAATC AGCGTGCGAG TGCCTGTTTC AAATCGTCAA
    101  TCAAATCGCC AACATATTCC AAACCGACCG ACAGGCGCAC CAATCCGGGG
    151  CGGATGTTGG CGGCGAGTTT TTCTTCGGGC TGCATCCTGC CGTGCGTGGT
    201  TGTCCACGGG TGGGTAATGG TCGAGCGCAC GTCACCGAGG TTGGCGGTGC
    251  GGGAAAAGAG TTCCACGCCG TCCACAACTT TCCACGCCGC TTCTTGATCG
    301  GCAACTTCAA AGCCGATGAC GATGCCGCCG CCGTTTTGCT GTTTGCGGAT
    351  AAGCGCCGCC TGAGGATGGT CGGACAATCC GGTGTAG
```

This corresponds to the amino acid sequence <SEQ ID 232; ORF 073>:

```
m073.pep
      1  MCMPYKIRVS DGICCPMPSE TRNQRASACF KSSIKSPTYS KPTDRRTNPG
     51  RMLAASFSSG CILPCVVVHG WVMVERTSPR LAVREKSSTP STTFHAASXS
    101  ATSKPMTMPP PFCCLRISAA XGWSDNPV*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 233>:

```
a073.seq
      1  ACGTGTATGT CATATAAGAT AAGGGTTTCA GACGGCATTT GCGGTGTTTT
     51  TCCGCCGATG CCGTCTGAA. CACGCAATCA GCGCGCGAGT GCCTGTTTCA
    101  AATCGTCAAT CAAATCGCCA ACATATTCCA AACCGACCGA CAGGCGCACC
    151  AATCCGGGGC GGATGTTGGC GGCGAGTTTT TCTTCGGGCT GCATCCTGCC
    201  GTGCGTGGTT GTCCACGGAT GGGTAATGGT CGAGCGCACG TCGCCGAGGT
    251  TGGCGGTACG GGAGAAAAGT TCGACGCCGT CCACGACTTT CCACGCGGCT
    301  GCTTGGTCGG CGACTTCAAA GCCGATGACG ATGCCGCCGC CGTTTTGCTG
    351  TTTGCGGATA AGCTCCGCCT GAGGATGGTC GGGTAATCCG GTGTAA
```

This corresponds to the amino acid sequence <SEQ ID 234; ORF 073.a>:

```
a073.pep
      1  TCMSYKIRVS DGICGVFPPM PSEXRNQRAS ACFKSSIKSP TYSKPTDRRT
     51  NPGRMLAASF SSGCILPCVV VHGWVMVERT SPRLAVREKS STPSTTFHAA
    101  AWSATSKPMT MPPPFCCLRI SSA*GWSGNP V*
```

m073/a073 92.3% identity over a 130 aa overlap

```
                 10         20         30         40         50
m073.pep    MCMPYKIRVSDGICC---PMPSETRNQRASACFKSSIKSPTYSKPTDRRTNPGRMLAASF
             || ||||||||||    |||||:||||||||||||||:|||||||||||||||||||||
a073        TCMSYKIRVSDGICGVFPPMPSEXRNQRASACFKSSIKSPTYSKPTDRRTNPGRMLAASF
                    10        20        30        40        50        60

            60        70        80        90       100       110
m073.pep    SSGCILPCVVVHGWVMVERTSPRLAVREKSSTPSTTFHAASXSATSKPMTMPPPFCCLRI
            |||||||||||||||||||||||||||||||||||||||||: |||||||||||||||||
a073        SSGCILPCVVVHGWVMVERTSPRLAVREKSSTPSTTFHAAAWSATSKPMTMPPPFCCLRI
                    70        80        90       100       110       120

            120       129
m073.pep    SAAXGWSDNPVX
            |:||||| ||||
a073        SSAXGWSGNPVX
                   130
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 073 shows 87.0% identity over a 131 aa overlap with a predicted ORF (ORF 073.ng) from *N. gonorrhoeae:*

m073/g073

```
                 10         20         30         40         50
m073.pep    MCMPYKIRVSDGICC---PMPSETRNQRASACFKSSIKSPTYSKPTDRRTNPGRMLAASF
            ||||| ||||||||    |||||||||||||||||||||||||||||||||:|||: ||||
g073        MCMPYAIRVSDGICRIFPPMPSETRNQRASACFKSSIKSPTYSKPTDRRTSPGRIPAASF
                    10        20        30        40        50        60

            60        70        80        90       100       110
m073.pep    SSGCILPCVVVHGWVMVERTSPRLAVREKSSTPSTTFHAASXSATSKPMTMPPPFCCLRI
            ||||||||||||| ||||||||||||||||||   |||||: |||||||||||||||||
g073        SSGCILPCVVVHGLVMVERTSPRLAVREKSST---TFHAAAWSATSKPMTMPPPFCCLRI
                    70        80        90       100       110

            120       129
m073.pep    SAAXGWSDNPVX
            |:| ||| ||||
g073        SSACGWSGNPVX
            120
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 235>:

g075.seq

```
  1  ATGCCGCCTT ACTTCATCAC CCTCTTAACG ATGGAAAATA CAAAAAGCGC
 51  GGCGAAAACG CCCACTACAA TCCAACCGGC TTCCATACCG TCCGCTTTTG
101  CGGCTTCCAA AGCGTTTTTT GCCGTTTCGG GCAACGCTGC GTTTGCCTGT
151  GCCGCCAAAG CCAGCGGGGC GGCTGTTACA ACAGCCAGTT TTGCGCCGTA
201  TTTACGGCAG GTGTTAATAA ATTTCATGAT ATTTTCCTTT ACGAAATTTT
251  TAAAAAAATG TGTTTGCGGG CTTTGTGAAG GTTTTAGAGA CCGCCTGCCG
301  GGCCTCTTAA ACTTAATCTT CTTTTTCGTA GAATCCGAAA ATTACAAATT
351  CCCCGCCTAT CTCTTCCAAT GCCGAGCTAA AAGCGTCTTC ATAGCTGTCA
401  TATTTACCGG CTGA
```

This corresponds to the amino acid sequence <SEQ ID 236; ORF 075.ng>:

g075.pep

```
  1  MPPYFITLLT MENTKSAAKT PTTIQPASIP SAFAASKAFF AVSGNAAFAC
 51  AAKASGAAVT TASFAPYLRQ VLINFMIFSF TKFLKKCVCG LCEGFRDRLP
101  GLLNLIFFFV ESENYKFPAY LFQCRAKSVF IAVIFTG*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 237>:

m075.seq
```
  1  ATGCCGTCTT ACTTCATCAC TCTCTTAACG ATGGAAAATA CAAAAAGCGC
 51  GGCGAAAATG CCCACTACAA TCCAACCGGC TTCCATACCG TCCGCTTTTG
101  CGGCTTCCAA AGCGTTTTTT GCCGTATCGG GCAACGTTGC ATTTGCATGT
151  GCGGCCAAAG CCAGGGGAGC AGCTGTTACA ACAGCCAGTT TTGCGCCGTA
201  TTTACGGCAG GTGTTAATAA ATTTCATGAT ATTTTCCTTC AAAAAGTGTT
251  TGGCGGTAAT GGATGGAGCG TTTTTCAGAC GACCGCCGAA CATCCGAAAA
301  TCAGTCTTTC AAAAATCCGA ATACGACAAA TTCGTATTGG TTGCCGATTT
351  CTTCCAAACC TGCGTTAATC GCTTCTTCGA AGTCGTAGAA ATAATCGGCA
401  TTGGTGATTA A
```

This corresponds to the amino acid sequence <SEQ ID 238; ORF 075>:

m075.pep
```
  1  MPSYFITLLT MENTKSAAKM PTTIQPASIP SAFAASKAFF AVSGNVAFAC
 51  AAKARGAAVT TASFAPYLRQ VLINFMIFSF KKCLAVMDGA FFRRPPNIRK
101  SVFQKSEYDK FVLVADFFQT CVNRFFEVVE IIGIGD*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 075 shows 65.7% identity over a 137 aa overlap with a predicted ORF (ORF 075.ng) from *N. gonorrhoeae:*

m075/g075
```
                   10        20        30        40        50        60
m075.pep    MPSYFITLLTMENTKSAAKMPTTIQPASIPSAFAASKAFFAVSGNVAFACAAKARGAAVT
            || ||||||||||||||||| ||||||||||||||||||||||||:|||||||| |||||
g075        MPPYFITLLTMENTKSAAKTPTTIQPASIPSAFAASKAFFAVSGNAAFACAAKASGAAVT
                   10        20        30        40        50        60

                   70        80            90       100       110
m075.pep    TASFAPYLRQVLINFMIFSF----KKCLAVMDGAFFRRPPNIRKSVFQKSEYDKFVLVAD
            ||||||||||||||||||||    |||:  :  :|  | |:: : :|   | ::: : |
g075        TASFAPYLRQVLINFMIFSFTKFLKKCVCGLCEGFRDRLPGLLNLIFFFVESENYKFPAY
                   70        80        90       100       110       120

          120       130
m075.pep    FFQTCVNRFFEVVEIIGIGDX
            :||  ::  | :| : |
g075        LFQCRAKSVFIAVIFTGX
                   130
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 239>:

a075.seq
```
  1  ATGCCGTCTT ACTTCATCAC TCTCTTAACG ATGGAAAAGA CAAAAAGCGC
 51  GGCGAAAACG CCCACTACAA TCCAACCGGC TTCCATACCG TCCGCTTTTG
101  CGGCTTCCAA AGCGTTTTTT GCTGTATCGG GCAACGTTGC ATTTGCATGT
151  GCGGCCAAAG CCAGGGGAGC AGCTGTTACA ACAGCCAGTT TTGCGCCGTA
201  TTTACGGCAG GTGTTAATAA ATTTCATGAT ATTTTCCTTC AAAAAGTGTT
251  TGGCGGTAAT GGATGGAGCG TTTTTCAGAC GACCGCCGAA CATCCGAAAA
301  TCAGTCTTTC AAAAATCCGA ATACGACAAA TTCGTATTGG TTGCCGATTT
351  CTTCCAAACC TGCGTTAATC GCTTCTTCGA AGTCGTAGAA ATAATCGGCA
401  TTGGTGATTA A
```

This corresponds to the amino acid sequence <SEQ ID 240; ORF 075.a>:

a075.pep
```
  1  MPSYFITLLT MEKTKSAAKT PTTIQPASIP SAFAASKAFF AVSGNVAFAC
 51  AAKARGAAVT TASFAPYLRQ VLINFMIFSF KKCLAVMDGA FFRRPPNIRK
101  SVFQKSEYDK FVLVADFFQT CVNRFFEVVE IIGIGD*
```

m075/a075 98.5% identity over a 136 aa overlap

```
                     10        20        30        40        50        60
m075.pep    MPSYFITLLTMENTKSAAKMPTTIQPASIPSAFAASKAFFAVSGNVAFACAAKARGAAVT
            ||||||||||||:|||||| ||||||||||||||||||||||||||||||||||||||||
a075        MPSYFITLLTMEKTKSAAKTPTTIQPASIPSAFAASKAFFAVSGNVAFACAAKARGAAVT
                     10        20        30        40        50        60

                     70        80        90       100       110       120
m075.pep    TASFAPYLRQVLINFMIFSFKKCLAVMDGAFFRRPPNIRKSVFQKSEYDKFVLVADFFQT
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a075        TASFAPYLRQVLINFMIFSFKKCLAVMDGAFFRRPPNIRKSVFQKSEYDKFVLVADFFQT
                     70        80        90       100       110       120

                    130
m075.pep    CVNRFFEVVEIIGIGDX
            ||||||||||||||||
a075        CVNRFFEVVEIIGIGDX
                    130
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 241>:

```
g080.seq
      1  ATGTGGGATA ATGCCGAAGC GATGGAACGG CTGACGCGCT GGCTGCTTGT
     51  CATGATGGCG ATGCTGCTTG CTGCGTCCGG GCTGGTTTGG TTTTACAATT
    101  CGAATCATCT GCCCGTCAAG CAGGTGTCGC TGAAGGGCAA CCTGGTTTAT
    151  TCCGATAAGA AGGCATTGGG CAGTTTGGCG AAAGAATACA TCCATGGGAA
    201  TATTTTGAGG ACGGACATCA ATGGCGCACA GGAAGCCTAC CGCCGGTATC
    251  CGTGGATTGC GTCGGTCATG GTGCGCCGCC GTTTTCCCGA TACGGTTGAG
    301  GTCGTCCTGA CCGAGCGCAA GCCGGTTGCA CGTTGGGGCG ACCATGCCTT
    351  GGTGGACGGC GAAGGCAATG TTTTTGAAGC CCGCTTGGAC AGACCCGGAA
    401  TGCCGGTATT CAGAGGCGCG GAAGGAACGT CTGCCGAAAT GCTCCGCCGT
    451  TATGACGAAT TTTCGACTGT TTTGGCAAAA CAGGGTTTGG GCATCAAAGA
    501  GATGACCTAT ACGGCACGTT CGGCGTGGAA TGTCGTTTTG GACAACGGCA
    551  TCACCGTCAG GCTCGGACGG GAAAAcgaGA TGAAACGCCT CCgGCTTTTT
    601  ACCgAAGCGT GGCAGCATCT gttgcGTAAG AATAAAAATC GGTTATCCTA
    651  TGTGGATATG Aggtataagg acggatttTC agtcccccat gctCCCGACG
    701  GTTTACCCGA AAAAGAATcc gAAGAATatt gggaacaggt ttgggacata
    751  ttacggcctg gcgtcggaaa cggttcgacg caaatttcaa tcagttatAA
    801  GGGCAGacga acaatggaac AGcagtaa
```

This corresponds to the amino acid sequence <SEQ ID 242; ORF 080.ng>:

```
g080.pep
      1  MWDNAEAMER LTRWLLVMMA MLLAASGLVW FYNSNHLPVK QVSLKGNLVY
     51  SDKKALGSLA KEYIHGNILR TDINGAQEAY RRYPWIASVM VRRRFPDTVE
    101  VVLTERKPVA RWGDHALVDG EGNVFEARLD RPGMPVFRGA EGTSAEMLRR
    151  YDEFSTVLAK QGLGIKEMTY TARSAWNVVL DNGITVRLGR ENEMKRLRLF
    201  TEAWQHLLRK NKNRLSYVDM RYKDGFSVPH APDGLPEKES EEYWEQVWDI
    251  LRPGVGNGST QISISYKGRR TMEQQ*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 243>:

```
m080.seq
      1  ATGTGGGATA ATGCCGAAGC GATGGAACGG CTGACGCGCT GGCTGCTTGT
     51  CATGATGGCG ATGCTGCTTG CTGCGTCCGG GCTGGTTTGG TTTTACAATT
    101  CGAATCATCT GCCCGTCAAG CAGGTGTCGC TGAAGGGCAA CCTGGTTTAT
    151  TCCGATAAGA AGACATTGGG CAGTTTGGCG AAAGAATACA TCCATGGGAA
    201  TATTTTGAGG ACGGACATCA ATGGCGCACA GGAGGCCTAC CGCCGGTATC
    251  CGTGGATTGC GTCGGTCATG GTGCGCCGCC GTTTTCCCGA CACGGTTGAG
    301  GTCGTCCTGA CCGAGCGCAA GCCGGTCGCG CGTTGGGGCG ACCATGCCTT
    351  GGTGGACGGC GAAGGCAATG TTTTTGAAGC CCGCTTGGAC AGACCCGGAA
    401  TGCCGGTATT CAGAGGCGCG GAAGGAACGT CTGCCGAAAT GCTCCGCCGT
    451  TATGACGAAT TTTCGACTGT TTTGGCAAAA CAGGGTTTGG GCATCAAAGA
    501  GATGACCTAT ACGGCACGTT CGGCGTGGAT TGTCGTTTTG GACAACGGCA
    551  TCACCGTCAG GCTCGGACGG GAAAACGAGA TGAAACGCCT CCGGCTTTTT
```

```
601  ACCGAAGCGT GGCAGCATCT GTTGCGTAAA AATAAAAATC GGTTATCCTA
651  TGTGGATATG AGGTATAAGG ACGGATTTTC AGTCCGCTAT GCTTCCGACG
701  GTTTACCCGA AAAAGAATCC GAAGAATAG
```

This corresponds to the amino acid sequence <SEQ ID 2441; ORF 080>:

```
m080.pep
  1  MWDNAEAMER LTRWLLVMMA MLLAASGLVW FYNSNHLPVK QVSLKGNLVY
 51  SDKKTLGSLA KEYIHGNILR TDINGAQEAY RRYPWIASVM VRRRFPDTVE
101  VVLTERKPVA RWGDHALVDG EGNVFEARLD RPGMPVFRGA EGTSAEMLRR
151  YDEFSTVLAK QGLGIKEMTY TARSAWIVVL DNGITVRLGR ENEMKRLRLF
201  TEAWQHLLRK NKNRLSYVDM RYKDGFSVRY ASDGLPEKES EE*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 080 shows 97.9% identity over a 242 aa overlap with a predicted ORF (ORF 080.ng) from *N. gonorrhoeae:*

```
m080/g080

                    10        20        30        40        50        60
m080.pep    MWDNAEAMERLTRWLLVMMAMLLAASGLVWFYNSNHLPVKQVSLKGNLVYSDKKTLGSLA
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
080         MWDNAEAMERLTRWLLVMMAMLLAASGLVWFYNSNHLPVKQVSLKGNLVYSDKKALGSLA
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m080.pep    KEYIHGNILRTDINGAQEAYRRYPWIASVMVRRRFPDTVEVVLTERKPVARWGDHALVDG
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
080         KEYIHGNILRTDINGAQEAYRRYPWIASVMVRRRFPDTVEVVLTERKPVARWGDHALVDG
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m080.pep    EGNVFEARLDRPGMPVFRGAEGTSAEMLRRYDEFSTVLAKQGLGIKEMTYTARSAWIVVL
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
080         EGNVFEARLDRPGMPVFRGAEGTSAEMLRRYDEFSTVLAKQGLGIKEMTYTARSAWNVVL
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m080.pep    DNGITVRLGRENEMKRLRLFTEAWQHLLRKNKNRLSYVDMRYKDGFSVRYASDGLPEKES
            |||||||||||||||||||||||||||||||||||||||||||||||||  :| |||||||
080         DNGITVRLGRENEMKRLRLFTEAWQHLLRKNKNRLSYVDMRYKDGFSVPHAPDGLPEKES
                   190       200       210       220       230       240

m080.pep    EEX
            ||
080         EEYWEQVWDILRPGVGNGSTQISISYKGRRTMEQQX
                   250       260       270
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 245>:

```
a080.seq
  1  ATGTGGGATA ATGCCGAAGC GATGGAACGG CTGACGCGCT GGCTGCTTGT
 51  CATGATGGCG ATGCTGCTTG CTGCGTCCGG GCTGGTTTGG TTTTACAATT
101  CGAATCATCT GCCCGTCAAG CAGGTGTCGC TGAAGGGCAA CCTAGTTTAT
151  TCCGATAAGA AAGCATTGGG CAGTTTGGCG AAAGAATACA TCCATGGGAA
201  TATTTTGAGG ACGGACATCA ATGGCGCACA GGAGGCCTAC CGCCGGTATC
251  CGTGGATTGC GTCGGTCATG GTGCGCCGCC GTTTTCCCGA CACGGTTGAG
301  GTCGTCCTGA CCGAGCGCAA GCCGGTCGCG CGTTGGGGCG ACCATGCCTT
351  GGTGGACGGC GAAGGCAATG TTTTTGAAGC CCGTTTGGAC AGACCCGGAA
401  TGCCGGTATT CAGAGGCGCG GAAGGAACGT CTGCCGAAAT GCTCCGCCGT
451  TATGACGAAT TTTCGACTGT TTTGGCAAAA CAGGGTTTGG GCATCAAAGA
501  GATGACCTAT ACGGCACGTT CGGCGTGGAT TGTCGTTTTG GACAACGGCA
```

```
551  TCACCGTCAG GCTCGGACGG GAAAACGAGA TGAAACGCCT CCGGCTTTTT
601  ACCGAAGCGT GGCAACATCT GTTGCGTAAA AATAAAAATC GGTTATCCTA
651  TGTGGATATG AGGTATAAGG ACGGATTTTC AGTCCGCTAT GCTCCCGACG
701  GTTTACCCGA AAAAGAATCC GAAGAATAG
```

This corresponds to the amino acid sequence <SEQ ID 246; ORF 080.a>:

a080.pep

```
  1  MWDNAEAMER LTRWLLVMMA MLLAASGLVW FYNSNHLPVK QVSLKGNLVY
 51  SDKKALGSLA KEYIHGNILR TDINGAQEAY RRYPWIASVM VRRRFPDTVE
101  VVLTERKPVA RWGDHALVDG EGNVFEARLD RPGMPVFRGA EGTSAEMLRR
151  YDEFSTVLAK QGLGIKEMTY TARSAWIVVL DNGITVRLGR ENEMKRLRLF
201  TEAWQHLLRK NKNRLSYVDM RYKDGFSVRY APDGLPEKES EE*
```

m080/a080 99.2% identity over a 242 aa overlap

```
                  10        20        30        40        50        60
m080.pep  MWDNAEAMERLTRWLLVMMAMLLAASGLVWFYNSNHLPVKQVSLKGNLVYSDKKTLGSLA
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
a080      MWDNAEAMERLTRWLLVMMAMLLAASGLVWFYNSNHLPVKQVSLKGNLVYSDKKALGSLA
                  10        20        30        40        50        60

                  70        80        90       100       110       120
m080.pep  KEYIHGNILRTDINGAQEAYRRYPWIASVMVRRRFPDTVEVVLTERKPVARWGDHALVDG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a080      KEYIHGNILRTDINGAQEAYRRYPWIASVMVRRRFPDTVEVVLTERKPVARWGDHALVDG
                  70        80        90       100       110       120

                 130       140       150       160       170       180
m080.pep  EGNVFEARLDRPGMPVFRGAEGTSAEMLRRYDEFSTVLAKQGLGIKEMTYTARSAWIVVL
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a080      EGNVFEARLDRPGMPVFRGAEGTSAEMLRRYDEFSTVLAKQGLGIKEMTYTARSAWIVVL
                 130       140       150       160       170       180

                 190       200       210       220       230       240
m080.pep  DNGITVRLGRENEMKRLRLFTEAWQHLLRKNKNRLSYVDMRYKDGFSVRYASDGLPEKES
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
a080      DNGITVRLGRENEMKRLRLFTEAWQHLLRKNKNRLSYVDMRYKDGFSVRYAPDGLPEKES
                 190       200       210       220       230       240

m080.pep  EEX
          |||
a080      EEX
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 247>:

g081.seq

```
  1  ATGAAACCAC TGGACCTAAA TTTCATCTGC CAAGCCCTCA AGCTTCCGAT
 51  GCCGTCTGAA AACAAACCCG TGTCGCGCAT CGTAACCGAC AGCCGCGATA
101  TTCGGGAAGG CGATGTGTTT TTCGCATTGG CGGGCGGGCG GTTTGACGCG
151  CATGATTTTG TTGGAGGCGT ATTGTCTGCG GGCGCGGCGG CGGTTGTGGT
201  TTCGCGCGAA GATTGCGCGG CTTTGGGCGG CGCGTTGAAA GTCGATGACA
251  CGCTTGCCGC GTTGCAAACG TTGGCGAAGG CGTGGCGCGA TAATGTGAAC
301  CCGTTTGTGT TCGGCATTAC CGGTTCGGGC GGCAAGACGA CGGTGAAGGA
351  GATGCTGGCT GCGGTATTGC GCCGCCGTTT CGGCGATGAT GCCGTTTCGG
401  CGACGGCAGG CAACTTCAAC AACCACAtcg gaTTGCCGCT GACTTTATTG
451  AAATtaaAcg aAAAACACCG CTATGCCGTG ATTGAAATGG GCATGAACCA
501  TTTTGGcgaa ctggcggtTt taacgcaaaT CGCCAAACCC GATGCCGCTT
551  TGGtcaACAA CGCCCTGCGC GCCCATGTCG GATGCGGTTt cgacggagtg
601  GGCGATATTG CCAAAGcgaa aagcGAGATT TatgcagGct tATGTTCAGA
651  CGGCATGGCA CTGATTCCTC AAGAAGATGC CAATATGGCT GTCTTCAAAA
701  CGGCAACGTT TAATTTGAAT ACGTGCACTT TCGGCGTCGA TAGCGGCGAT
751  GTCCGCGCGG AAAATATCGT GCTGAAACCT TTGTCGTGCG AATTTGATTT
```

801 GGTGTGCGGC GACGAGCGCA CTGCCGTGGT GCTGCCTGTT CCCGGCCGCC
851 ACAATGTCCA CAACGCCGCC GCTGCCGCCG CGCTGGCTTT GGCTGCCGGT
901 TTGAGTTTGA ACGATGTGGC GGAAGGTTTG CAAGGCTTCA GCAACATCAA
951 AGGCCGTCTG AACGTCAAAG CCGGCATCAA GGGCGCAACC CTGATTGACG
1001 ATACTTATAA TGCGAATCCC GACAGTATGA AAGCCGCGGT TGACGTGTTG
1051 GCGCGTATGC CTGCGCCGCG CATTTTCGTG ATGGGCGATA TGGGCGAACT
1101 GGGCGAGGAc gaAGCCGCCG CCATGCACGC CGAagtcgGC GCGTACGCCC
1151 GCGACCAAGG CATCGAAGCG GCTTATTTTG TCGGCGACAA CAGCGTCGAA
1201 GCGGcggaAA AATTTGGCGC GGACGGTTTG TGGTTCGCCG CCAAAGACCC
1251 GTTGATTCAA GTGTTGAGCC ACGATTTGCC CGAACGCGCC ACCGTGTTGG
1301 TGAAAGGTTC GCGCTTTATG CAGAtggAAG AAGTGGTCGA GGCATTGGAG
1351 GATAAGTga

This corresponds to the amino acid sequence <SEQ ID 248; ORF 081.ng>:

g081.pep

1 MKPLDLNFIC QALKLPMPSE NKPVSRIVTD SRDIREGDVF FALAGGRFDA
51 HDFVGGVLSA GAAAVVVSRE DCAALGGALK VDDTLAALQT LAKAWRDNVN
101 PFVFGITGSG GKTTVKEMLA AVLRRRFGDD AVSATAGNFN NHIGLPLTLL
151 KLNEKHRYAV IEMGMNHFGE LAVLTQIAKP DAALVNNALR AHVGCGFDGV
201 GDIAKAKSEI YAGLCSDGMA LIPQEDANMA VFKTATFNLN TCTFGVDSGD
251 VRAENIVLKP LSCEFDLVCG DERTAVVLPV PGRHNVHNAA AAAALALAAG
301 LSLNDVAEGL QGFSNIKGRL NVKAGIKGAT LIDDTYNANP DSMKAAVDVL
351 ARMPAPRIFV MGDMGELGED EAAAMHAEVG AYARDQGIEA AYFVGDNSVE
401 AAEKFGADGL WFAAKDPLIQ VLSHDLPERA TVLVKGSRFM QMEEVVEALE
451 DK*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 249>:

m081.seq

1 ATGAAACCAC TGGACCTAAA TTTCATCTGC CAAGCCCTCA AGCTTCCGAT
51 GCCGTCTGAA AGCAAACCCG TGTCGCGCAT CGTAACCGAC AGCCGCGACA
101 TCCGCGCGGG CGATGTGTTT TTCGCATTGG CGGGCGAGCG GTTTGACGCG
151 CATGATTTTG TTGAAGACGT ATTGGCTGCT GGTGCGGCGG CGGTTGTGGT
201 TTCGCGCGAA GATTGTGCTG CAATGGATGG CGCGTTGAAA GTCGATGACA
251 CGCTTGCCGC ATTGCAAACG CTGGCAAAGG CGTGGCGTGA AAATGTGAAT
301 CCGTTTGTGT TCGGCATTAC CGGTTCGGGC GGCAAGACGA CGGTGAAGGA
351 AATGCTGGCT GCGGTATTGC GCCgCCGTTT CGGCGATGAT GCCGTGTTGG
401 CGACGGCAGG CAACTTCAAC AACCATATCG GATTGCCGCT GACTTTGTTG
451 AAGTTAAACG AAAAACACCG CTATGCCGTG ATTGAAATGG GCATGAACCA
501 TTTCGGCGAA CTGGCGGTTT TAACGCAmAT CGCCAAACCA AATGCCGCAT
551 TGGTCAACAA CGCCATGCGC GCCCATGTCG GCTGCGGTTT CGACGGAGTG
601 GGCGATATTG CCAAAGCGAA AAGCGAGATT TACCAAGGTT TATGTTCAGA
651 CGGCATTGCA CTGATTCCTC AAGAAGATGC CAATATGGCT GTCTTCAAAA
701 CGGCAACGCT TAATTTGAAT ACGCGCACTT TCGGCATCGA TAGCGGCGAT
751 GTTCACGCGG AAAATATTGT GCTGAAACCG TTGTCGTGCG AATTTGATTT
801 GGTGTGCGGC GATGAGCGCG CCGCCGTGGT GCTGCCTGTT CCCGGCCGCC
851 ACAATGTCCA CAACGCCGCC GCTGCCGCCG CGCTGGCTTT GGCTGCGGGT
901 TTGAGTTTGA ACGATGTGGC GGAAGGTTTG AAAGGCTTCA GCAATATCAA
951 AGGCCGTCTG AACGTCAAAT CCGGAATCAA GGGCGCAACC CTGATTGACG
1001 ATACTTATAA TGCGAACCCT GACAGCATGA AAGCTGCGAT TGACGTGTTG
1051 GCGCGTATGC CTGCGCCGCG TATTTTCGTG ATGGGCGATA TGGGCGAACT
1101 GGGCGAACTG GGCGAGGACG AAGCCGCCGC TATGCACGCC GAAGTCGGCG
1151 CGTATGCCCG CGACCAAGGC ATCGAAGCGG CTTATTTTGT CGGCGACAAC
1201 AGCGTCGAAG CGGCGGAAAA ATTTGGCGCG GACGGTTTGT GGTTCGCCGC
1251 CAAAGACCCG TTGATTCAAG TGTTGCGCCA CGATTTGCCC GAACGCGCCA
1301 CCGTGTTGGT GAAAGGTTCG CGCTTTATGC AGATGGAAGA AGTGGTCGAG
1351 GCATTGGAGG ATAAGTGA

This corresponds to the amino acid sequence <SEQ ID 250; ORF 081>:

m081.pep

1 MKPLDLNFIC QALKLPMPSE SKPVSF[illegible]TD SRDIRAGDVF FALAGERFDA
51 HDFVEDVLAA GAAAVVVSRE DCAAMDGALK VDDTLAALQT LAKAWRENVN
101 PFVFGITGSG GKTTVKEMLA AVLRRRFGDD AVLATAGNFN NHIGLPLTLL
151 KLNEKHRYAV IEMGMNHFGE LAVLTXIAKP NAALVNNAMR AHVGCGFDGV

```
201  GDIAKAKSEI YQGLCSDGIA LIPQEDANMA VFKTATLNLN TRTFGIDSGD
251  VHAENIVLKP LSCEFDLVCG DERAAVVLPV PGRHNVHNAA AAAALALAAG
301  LSLNDVAEGL KGFSNIKGRL NVKSGIKGAT LIDDTYNANP DSMKAAIDVL
351  ARMPAPRIFV MGDMGELGEL GEDEAAAMHA EVGAYARDQG IEAAYFVGDN
401  SVEAAEKFGA DGLWFAAKDP LIQVLRHDLP ERATVLVKGS RFMQMEEVVE
451  ALEDK*
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

ORF 081 shows 94.1% identity over a 455 aa overlap with a predicted ORF (ORF 081.ng) from *N. gonorrhoeae:*

m081/g081

```
                      10        20        30        40        50        60
m081.pep      MKPLDLNFICQALKLPMPSESKPVSRIVTDSRDIRAGDVFFALAGERFDAHDFVEDVLAA
              ||||||||||||||||||||:|||||||||||||| ||||||||| ||||||||: ||:|
g081          MKPLDLNFICQALKLPMPSENKPVSRIVTDSRDIREGDVFFALAGGRFDAHDFVGGVLSA
                      10        20        30        40        50        60

                      70        80        90       100       110       120
m081.pep      GAAAVVVSREDCAAMDGALKVDDTLAALQTLAKAWRENVNPFVFGITGSGGKTTVKEMLA
              ||||||||||||||: ||||||||||||||||||||:|||||||||||||||||||||||
g081          GAAAVVVSREDCAALGGALKVDDTLAALQTLAKAWRDNVNPFVFGITGSGGKTTVKEMLA
                      70        80        90       100       110       120

                     130       140       150       160       170       180
m081.pep      AVLRRRFGDDAVLATAGNFNNHIGLPLTLLKLNEKHRYAVIEMGMNHFGELAVLTXIAKP
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
g081          AVLRRRFGDDAVSATAGNFNNHIGLPLTLLKLNEKHRYAVIEMGMNHFGELAVLTQIAKP
                     130       140       150       160       170       180

                     190       200       210       220       230       240
m081.pep      NAALVNNAMRAHVGCGFDGVGDIAKAKSEIYQGLCSDGIALIPQEDANMAVFKTATLNLN
              :|||||||:|||||||||||||||||||||| ||||||:|||||||||||||||||:|||
g081          DAALVNNALRAHVGCGFDGVGDIAKAKSEIYAGLCSDGMALIPQEDANMAVFKTATFNLN
                     190       200       210       220       230       240

                     250       260       270       280       290       300
m081.pep      TRTFGIDSGDVHAENIVLKPLSCEFDLVCGDERAAVVLPVPGRHNVHNAAAAAALALAAG
              | |||:|||||:|||||||||||||||||||||:||||||||||||||||||||||||||
g081          TCTFGVDSGDVRAENIVLKPLSCEFDLVCGDERTAVVLPVPGRHNVHNAAAAAALALAAG
                     250       260       270       280       290       300

                     310       320       330       340       350       360
m081.pep      LSLNDVAEGLKGFSNIKGRLNVKSGIKGATLIDDTYNANPDSMKAAIDVLARMPAPRIFV
              ||||||||||:||||||||||||:||||||||||||||||||||||:|||||||||||||
g081          LSLNDVAEGLQGFSNIKGRLNVKAGIKGATLIDDTYNANPDSMKAAVDVLARMPAPRIFV
                     310       320       330       340       350       360

                     370       380       390       400       410       420
m081.pep      MGDMGELGELGEDEAAAMHAEVGAYARDQGIEAAYFVGDNSVEAAEKFGADGLWFAAKDP
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
g081          MGDMGELGE---DEAAAMHAEVGAYARDQGIEAAYFVGDNSVEAAEKFGADGLWFAAKDP
                          370       380       390       400       410

                     430       440       450
m081.pep      LIQVLRHDLPERATVLVKGSRFMQMEEVVEALEDKX
              ||||| ||||||||||||||||||||||||||||||
g081          LIQVLSHDLPERATVLVKGSRFMQMEEVVEALEDKX
              420       430       440       450
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 251>:

a081.seq

```
   1  ATGAAACCAC TGGACCTAAA TTTCATCTGC CAAGCCCTCA AGCTTCCGAT
  51  GCCGTCTGAA AGCAAACCCG TGTCGCGCAT CGTAACCGAC AGCCGCGACA
 101  TCCGCGCGGG CGATGTGTTT TTCGCATTGG CGGGCGGGCG GTTTGATGCG
 151  CATGATTTTG TTGAAGACGT ATTGGCTGCG GGTGCGGCGG CGGTTGTGGT
 201  TTCGCGCGAA GATTGCGTTG CAATGGATGG CGCGTTGAAA GTCGATGACA
 251  CGCTTACCGC GTTGCAAATG TTGGCGAAGG CGTGGCGCGA GAATGTGAAC
 301  CCGTTTGTGT TCGGTATTAC CGGCTCGGGC GGCAAGACCA CGGTGAAGGA
 351  AATGTTGGCT GCGGTATTGC GCCGCCGTTT CGGCGATAAT GCCGTTTTGG
 401  CGACGGCAGG CAACTTCAAC AACCACATCG GATTGCCGTT GACTTTGTTG
 451  AAATTAAACG AAAAACACCG CTATGCCGTG ATTGAAATGG GTATGAACCA
 501  TTTTGGCGAA CTGGCGGTTT TGACACAAAT CGCCAAACCC GATGCCGCAT
 551  TGGTCAACAA CGCCATGCGC GCCCATGTCG GCTGCGGTTT CGACGGAGTG
 601  GGCGATATTG CCAAAGCGAA AAGCGAGATT TATCAAGGCT TATGTTCAGA
 651  CGGCATGGCG CTGATTCCTC AAGAAGATGC CAATATGGCT GTCTTCAAAA
 701  CGGCAACGCT TAATTTGAAT ACGCGCACTT TCGGCATCGA TAGCGGCGAT
 751  GTCCACGCGG AAAATATCGT GCTGAAACCG TTGTCGTGCG AATTTGATTT
 801  GGTGTGCGGC AACGAGTGCG CAGCCGTGGT TCTGCCCGTT CCCGGCCGCC
 851  ACAATGTCCA CAACGCCGCC GCCGCCGCCG CGCTGTCTTT GGCTGCAGGT
 901  TTGAGTTTGA ACGATGTGGC GGAAGGTTTG AAAGGCTTCA GCAATATCAA
 951  AGGCCGTCTG AACGTCAAAT CCGGAATCAA GGGCGCAACC CTGATTGACG
1001  ATACTTATAA TGCGAACCCT GACAGCATGA AAGCTGCGGT TGACGTGTTG
1051  GCGCGTATGC CTGCGCCGCG TATTTTCGTG ATGGGCGATA TGGGCGAACT
1101  GGGTGAGGAC GAAGCCGCCG CCATGCACGC CGAAGTCGGC GCGTACGCCC
1151  GCGACCAAGG CATCGAAGCG GCTTATTTTG TCGGCGACAA CAGCGTCGAA
1201  GCGGCGGAAA AATTTGGCGC GGACGGTTTG TGGTTCGCCG CCAAAGACCC
1251  GTTGATTCAA GTGTTGCGCC ACGATTTGCC CGAACGCGCC ACCGTGTTGG
1301  TGAAAGGTTC GCGCTTTATG CAGATGGAAG AAGTGGTCGA GGCATTGGAG
1351  GATAAGTGA
```

This corresponds to the amino acid sequence <SEQ ID 252; ORF 081.a>:

a081.pep

```
   1  MKPLDLNFIC QALKLPMPSE SKPVSRIVTD SRDIRAGDVF FALAGGRFDA
  51  HDFVEDVLAA GAAAVVVSRE DCVAMDGALK VDDTLTALQM LAKAWRENVN
 101  PFVFGITGSG GKTTVKEMLA AVLRRRFGDN AVLATAGNFN NHIGLPLTLL
 151  KLNEKHRYAV IEMGMNHFGE LAVLTQIAKP DAALVNNAMR AHVGCGFDGV
 201  GDIAKAKSEI YQGLCSDGMA LIPQEDANMA VFKTATLNLN TRTFGIDSGD
 251  VHAENIVLKP LSCEFDLVCG NECAAVVLPV PGRHNVHNAA AAALSLAAG
 301  LSLNDVAEGL KGFSNIKGRL NVKSGIKGAT LIDDTYNANP DSMKAAVDVL
 351  ARMPAPRIFV MGDMGELGED EAAAMHAEVG AYARDQGIEA AYFVGDNSVE
 401  AAEKFGADGL WFAAKDPLIQ VLRHDLPERA TVLVKGSRFM QMEEVVEALE
 451  DK*
```

m081/a081 96.7% identity over a 455 aa overlap

```
                    10        20        30        40        50        60
m081.pep    MKPLDLNFICQALKLPMPSESKPVSRIVTDSRDIRAGDVFFALAGERFDAHDFVEDVLAA
            |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
a081        MKPLDLNFICQALKLPMPSESKPVSRIVTDSRDIRAGDVFFALAGGRFDAHDFVEDVLAA
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m081.pep    GAAAVVVSREDCAAMDGALKVDDTLAALQTLAKAWRENVNPFVFGITGSGGKTTVKEMLA
            |||||||||||:|||||||||||||:||| ||||||||||||||||||||||||||||||
a081        GAAAVVVSREDCVAMDGALKVDDTLTALQMLAKAWRENVNPFVFGITGSGGKTTVKEMLA
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m081.pep    AVLRRRFGDDAVLATAGNFNNHIGLPLTLLKLNEKHRYAVIEMGMNHFGELAVLTXIAKP
            |||||||||:|||||||||||||||||||||||||||||||||||||||||||||| ||||
a081        AVLRRRFGDNAVLATAGNFNNHIGLPLTLLKLNEKHRYAVIEMGMNHFGELAVLTQIAKP
```

```
                         130       140       150       160       170       180

                         190       200       210       220       230       240
m081.pep         NAALVNNAMRAHVGCGFDGVGDIAKAKSEIYQGLCSDGIALIPQEDANMAVFKTATLNLN
                 :|||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
a081             DAALVNNAMRAHVGCGFDGVGDIAKAKSEIYQGLCSDGMALIPQEDANMAVFKTATLNLN
                         190       200       210       220       230       240

                         250       260       270       280       290       300
m081.pep         TRTFGIDSGDVHAENIVLKPLSCEFDLVCGDERAAVVLPVPGRHNVHNAAAAAALALAAG
                 ||||||||||||||||||||||||||||||:| ||||||||||||||||||||||:||||
a081             TRTFGIDSGDVHAENIVLKPLSCEFDLVCGNECAAVVLPVPGRHNVHNAAAAAALSLAAG
                         250       260       270       280       290       300

                         310       320       330       340       350       360
m081.pep         LSLNDVAEGLKGFSNIKGRLNVKSGIKGATLIDDTYNANPDSMKAAIDVLARMPAPRIFV
                 ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
a081             LSLNDVAEGLKGFSNIKGRLNVKSGIKGATLIDDTYNANPDSMKAAVDVLARMPAPRIFV
                         310       320       330       340       350       360

                         370       380       390       400       410       420
m081.pep         MGDMGELGELGEDEAAAMHAEVGAYARDQGIEAAYFVGDNSVEAAEKFGADGLWFAAKDP
                 |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
a081             MGDMGELGE---DEAAAMHAEVGAYARDQGIEAAYFVGDNSVEAAEKFGADGLWFAAKDP
                            370       380       390       400       410

                         430       440       450
m081.pep         LIQVLRHDLPERATVLVKGSRFMQMEEVVEALEDKX
                 ||||||||||||||||||||||||||||||||||||
a081             LIQVLRHDLPERATVLVKGSRFMQMEEVVEALEDKX
                 420       430       440       450
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 253>:

```
g082.seq
      1  aTGTGGTTGT TGAAGTTGCC TGCCGTCGCC GAAACGGCAT CATCGCCGAA
     51  ACGGCGGCGC AATACCGCAG CCAGCATCTC CTTCACCGTC GTCTTGCCGC
    101  CCGAACCGGT AATGCCGAAC ACAAACGGGT TCACATTATC GCGCCACGCC
    151  TTCGCCAACG TTTGCAACGC GGCAAGCGTG TCATCGACTT TCAACGCGCC
    201  GCCCAAAGCC GCGCAATCTT CGCGCGAAAC CACAACCGCC GCCGCGCCCG
    251  CAGACAATAC GCCTCCAACA AAATCATGCG CGTCAAACCG CCCGCCCGCC
    301  AATGCGAAAA ACACATCGCC TTCCCGAATA TCGCGGCTGT CGGTTACGAT
    351  GCGCGACACG GGTTTGTTTT CAGACGGCAT CGGAAGCTTG AGGGCTTGGC
    401  AGATGAAATT TAGGTCCAGT GGTTTCATAT TTGCTTTCGT TAATATTCGG
    451  GCGGCGGACA CATCGGTAGC GGCTGATTTT TTTATCGCCT GTTTTGCTGT
    501  GGTAAAACAC AGATTATTTT CCCATTCTCA TTCGGCATTT TTTCTGTACG
    551  TATCATTTTT TAGACGTATT TTTAGCCGAT TTGCCTTTTC CCGCATACCA
    601  CGGCGCGGGG TCGTCGGACT GTCTGTCGAT AAAGGCAAGG TTATTGCCTT
    651  CGCCCGGCAC ATCGGGGACA TTCCCCCAAA AATCATAGCC GTCATCGGGC
    701  AACTCGTCGG TTTCGATACC CGTCCAACTG CCGAATCCGC GTAA
```

This corresponds to the amino acid sequence <SEQ ID 254; ORF 082.ng>:

```
g082.pep
      1  MWLLKLPAVA ETASSPKRRR NTAASISFTV VLPPEPVMPN TNGFTLSRHA
     51  FANVCNAASV SSTFNAPPKA AQSSRETTTA AAPADNTPPT KSCASNRPPA
    101  NAKNTSPSRI SRLSVTMRDT GLFSDGIGSL RAWQMKFRSS GFIFAFVNIR
    151  AADTSVAADF FIACFAVVKH RLFSHSHSAF FLYVSFFRRI FSRFAFSRIP
    201  RRGVVGLSVD KGKVIAFARH IGDIPPKIIA VIGQLVGFDT RPTAESA*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 255>:

```
m082.seq
      1  ATGnnGTTGT TGAAGTTGCC TGCCGTCGCC AACACGGCAT CATCGCCGAA
     51  ACGGcGGCGC AATACCGCAG CCAGCATTTC CTTCACCGTC GTCTTGCCGC
    101  CCGAACCGGT AATGCCGAAC ACAAACGGAT TCACATTTTC ACGCCACGCC
    151  TTTGCCAGCG TTTGCAATGC GGCAAGCGTG TCATCGACTT TCAACGCGCC
```

```
201  ATCCATTGCA GCACAATCTT CGCGCGAAAC CACAACCGCC GCCGCACCAG
251  CAGCCAATAC GTCTTCAACA AAATCATGCG CGTCAAACCG CTCGCCCGCC
301  AATGCGAAAA ACACATCGCC CGCGCGGATG TCGCGGCTGT CGGTTACGAT
351  GCGCGACACG GGTTTGCTTT CAGACGGCAT CGGAAGCTTG AGGGCTTGGC
401  AGATGAAATT TAGGTCCAGT GGTTTCATAT TTACTTTCGT TAATATTCGG
451  GCGGCGGACA CATCGGTAGC GGCTGATTTT TTTATCGCCT GTTTTGCTGT
501  GGTAAAACAC AGATTATTTT CCCATTCTCA TTCGGsATTT TTTCTGTACG
551  TATCATTTTT TAGACGTATT TTTAGTCGAT TTGCCTTTTC CCGCATACCA
601  CGGCGCGGGG TCGTCGGGCA GTCCGTCGAT AAAGGCAAGG TTATTGCCTT
651  CGCCCTGCAC ATCGGGAACA TTCCCCCAAA AATCATAGCC GTCATCGGGC
701  AACTCGTCGG TTTCGATACC CGTCCAACTG CCGAATCCGC GTAA
```

This corresponds to the amino acid sequence <SEQ ID 256: ORF 082>:

```
m082.pep
  1  MXLLKLPAVA NTASSPKRRR NTAASISFTV VLPPEPVMPN TNGFTFSRHA
 51  FASVCNAASV SSTFNAPSIA AQSSRETTTA AAPAANTSST KSCASNRSPA
101  NAKNTSPARM SRLSVTMRDT GLLSDGIGSL RAWQMKFRSS GFIFTFVNIR
151  AADTSVAADF FIACFAVVKH RLFSHSHSXF FLYVSFFRRI FSRFAFSRIP
201  RRGVVGQSVD KGKVIAFALH IGNIPPKIIA VIGQLVGFDT RPTAESA*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 082 shows 92.7% identity over a 247 aa overlap with a predicted ORF (ORF 082.ng) from *N. gonorrhoeae:*

```
m082/g082

                    10        20        30        40        50        60
m082.pep    MXLLKLPAVANTASSPKRRRNTAASISFTVVLPPEPVMPNTNGFTFSRHAFASVCNAASV
            | |||||||:||||||||||||||||||||||||||||||||||||:|||||:||||||
g082        MWLLKLPAVAETASSPKRRRNTAASISFTVVLPPEPVMPNTNGFTLSRHAFANVCNAASV
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m082.pep    SSTFNAPSIAAQSSRETTTAAAPAANTSSTKSCASNRSPANAKNTSPARMSRLSVTMRDT
            ||||||  |||||||||||||| ||  |||||||| ||||||||:|:||||||||||
g082        SSTFNAPPKAAQSSRETTTAAAPADNTPPTKSCASNRPPANAKNTSPSRISRLSVTMRDT
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m082.pep    GLLSDGIGSLRAWQMKFRSSGFIFTFVNIRAADTSVAADFFIACFAVVKHRLFSHSHSXF
            ||:|||||||||||||||||||||:||||||||||||||||||||||||||||||||| |
g082        GLFSDGIGSLRAWQMKFRSSGFIFAFVNIRAADTSVAADFFIACFAVVKHRLFSHSHSAF
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m082.pep    FLYVSFFRRIFSRFAFSRIPRRGVVGQSVDKGKVIAFALHIGNIPPKIIAVIGQLVGFDT
            |||||||||||||||||||||||||| ||||||||||| |||:|||||||||||||||||
g082        FLYVSFFRRIFSRFAFSRIPRRGVVGLSVDKGKVIAFARHIGDIPPKIIAVIGQLVGFDT
                   190       200       210       220       230       240

m082.pep    RPTAESAX
            ||||||||
g082        RPTAESAX
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 257>:

```
a082.seq
  1  ATGTGGTTGT TGAAGTTGCC TGCCGTCGCC AAAACGGCAT TATCGCCGAA
 51  ACGGCGGCGC AATACCGCAG CCAACATTTC CTTCACCGTC GTCTTGCCGC
```

```
101  CCGAGCCGGT AATACCGAAC ACAAACGGGT TCACATTCTC GCGCCACGCC
151  TTCGCCAACA TTTGCAACGC GGTAAGCGTG TCATCGACTT TCAACGCGCC
201  ATCCATTGCA ACGCAATCTT CGCGCGAAAC CACAACCGCC GCCGCACCCG
251  CAGCCAATAC GTCTTCAACA AAATCATGCG CATCAAACCG CCCGCCCGCC
301  AATGCGAAAA ACACATCGCC CGCGCGGATG TCGCGGCTGT CGGTTACGAT
351  GCGCGACACG GGTTTGCTTT CAGACGGCAT CGGAAGCTTG AGGGCTTGGC
401  AGATGAAATT TAGGTCCAGT GGTTTCATAT TTACTTTCGT TAATATTCGG
451  GCGGCGGACA CATCGGTAGC GGCTGATTTT TTTATCGCCT GTTTTGCTGT
501  GGTAAAACAC AGATTATTTT CCCATTCTCA TTCGGCATTT TTTCTGTACG
551  TATCATTTTT TAGACGTATT TTTAGTCGAT TTGCCTTTTC CCGCATACCA
601  CGGCGCGGGG TCGTCGGGCA GTCCGTCGAT AAAGGCAAGG TTATTGCCTT
651  CGCCCTGCAC ATCGGGAACA TTCCCCCAAA AATCATAGCC GTCATCGGGC
701  AACTCGTCGG TTTCGATACC CGTCCAACTG CCGAATCCGC GTAA
```

This corresponds to the amino acid sequence <SEQ ID 258; ORF 082.a>:

a082.pep

```
  1  MWLLKLPAVA KTALSPKRRR NTAANISFTV VLPPEPVIPN TNGFTFSRHA
 51  FANICNAVSV SSTFNAPSIA TQSSRETTTA AAPAANTSST KSCASNRPPA
101  NAKNTSPARM SRLSVTMRDT GLLSDGIGSL RAWQMKFRSS GFIFTFVNIR
151  AADTSVAADF FIACFAVVKH RLFSHSHSAF FLYVSFFRRI FSRFAFSRIP
201  RRGVVGQSVD KGKVIAFALH IGNIPPKIIA VIGQLVGFDT RPTAESA*
```

m082/a082 95.5% identity over a 247 aa overlap

```
                    10        20        30        40        50        60
m082.pep    MXLLKLPAVANTASSPKRRRNTAASISFTVVLPPEPVMPNTNGFTFSRHAFASVCNAASV
            | ||||||||:|| ||||||||||:||||||||||||:||||||||||||||::|||:||
a082        MWLLKLPAVAKTALSPKRRRNTAANISFTVVLPPEPVIPNTNGFTFSRHAFANICNAVSV
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m082.pep    SSTFNAPSIAAQSSRETTTAAAPAANTSSTKSCASNRSPANAKNTSPARMSRLSVTMRDT
            ||||||||||:||||||||||||||||||||||||||| ||||||||||||||||||||||
a082        SSTFNAPSIATQSSRETTTAAAPAANTSSTKSCASNRPPANAKNTSPARMSRLSVTMRDT
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m082.pep    GLLSDGIGSLRAWQMKFRSSGFIFTFVNIRAADTSVAADFFIACFAVVKHRLFSHSHSXF
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
a082        GLLSDGIGSLRAWQMKFRSSGFIFTFVNIRAADTSVAADFFIACFAVVKHRLFSHSHSAF
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m082.pep    FLYVSFFRRIFSRFAFSRIPRRGVVGQSVDKGKVIAFALHIGNIPPKIIAVIGQLVGFDT
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a082        FLYVSFFRRIFSRFAFSRIPRRGVVGQSVDKGKVIAFALHIGNIPPKIIAVIGQLVGFDT
                   190       200       210       220       230       240

m082.pep    RPTAESAX
            ||||||||
a082        RPTAESAX
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 259>:

g084.seq

```
  1  ATGAAacaAT CCGcccgaat aAAAAATATG GATCAGACAT TAAAAAATAc
 51  attgggcatt tGCGCGcttt tagcctTTTG TTTTggcgcG gccaTCGCAT
101  CAGGTTATCA CTTGGAATAT GAATACGGCT ACCGTTATTC TGCCGTGGGC
151  GCTTTGGCTT CGGTTGTATT TTTATTATTA TTGGCACGCG GCTTCCCGCG
201  CGTTTCTTCA GTTGTTTTAC TGATTTACGT CGGCACAACC GCCCTATATT
251  TGCCGGTCGG CTGGCTGTAT GGTGCGCCTT CTTATCAGAT AGTCGGTTCG
301  ATATTGGAAA GCAATCCTGC CGAGGCGCGT GAATTTGTCG GCAATCTTCC
351  CGGGTCGCTT TATTTTGTGC AGGCATTATT TTTCATTTTT GGCTTGACAG
```

```
401  TTTGGAAATA TTGTGTATCT GTGGGGGTAT TTGCTGACGT AAAAAACTAT
451  AAACGTCGCA GCAAAATATG GCTGACCATA TTATTGACTT TGATTTTGTC
501  CTGCGCGGTG ATGGAGAAAA TCGccggcga taaAGATTGG CGAGaacctg
551  atgccggcct gttgttgaat ATTTTcgacc tgtattaCga cttggctttc
601  cgcgccggca cAATATGCCG CCAAGCGCGC CCAcatttg gaagCagcaa
651  aaaaagcgtC AACATGGCAt atccgccaac ttgcgcccaa gTAtaa
```

This corresponds to the amino acid sequence <SEQ ID 260; ORF 084.ng>:

```
g084.pep
  1  MKQSARIKNM DQTLKNTLGI CALLAFCFGA AIASGYHLEY EYGYRYSAVG
 51  ALASVVFLLL LARGFPRVSS VVLLIYVGTT ALYLPVGWLY GAPSYQIVGS
101  ILESNPAEAR EFVGNLPGSL YFVQALFFIF GLTVWKYCVS VGVFADVKNY
151  KRRSKIWLTI LLTLILSCAV MEKIAGDKDW REPDAGLLLN IFDLYYDLAF
201  RAGTICRQAR PHFGSSKKSV NMAYPPTCAQ V*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 261>:

```
m084.seq
  1  ATGAAACAAT CCGCCcGAAT AAAa.ATATG AATCAGACAT TACTTTATAC
 51  ATTGGGCATT TGCGCGCTTT TAACCTTTnn nnnnnnnnnn nnnnnnnnnn
101  nnnnnTATCA CCCnGAATAT GAATACGGCT ACCGTTATTC TGCCGTGGGT
151  GCTTTGGCTT CGGTTGTATT TTTATTATTA TTGGCACGCG GTTTCCCGCG
201  CGTTTCTTCA GTTGTTTTAC TGATTTACGT CGGCACAACC GCCCTATATT
251  TGCCGGTCGG CTGGCTGTAT GGTGCGCCGT CTTATCAGAT AGTCGGTTCG
301  ATATTGGAAA GCAATCCTGC CGAGGCGCGT GAATTTGTCG GCAATCTTCC
351  CGGGTCGCTT TATTTTGTGC AGGCATTATT TTTCATTTTT GGCTTGACAG
401  TTTGGAAATA TTGTGTATCG GGGGGGGTAT TTGCTGACGT AAAAAACTAT
451  AAACGCCGCA GCAAAATATG GCTGACTATA TTATTGACTT TGATTTTGTC
501  CTGCGCGGTG ATGGATAAAA TCGCCAGCGA TAAAGATTTG CGAGAACCTG
551  ATGCCGGCCT GTTGTTGAAT ATTTTCGACC TGTATTACGA TTTGGCT.TC
601  CGCGCCGGCA CAATATGCCG CCAAGCGCGC CCACATTTTG GAAGCAGCAA
651  AAAAAGCGTC AACATGGCAT ATCCGTCATG TTGCGCCCAA GTATAA
```

This corresponds to the amino acid sequence <SEQ ID 262; ORF 084>:

```
m084.pep
  1  MKQSARIKXM NQTLLYTLGI CALLTFXXXX XXXXXYHPEY EYGYRYSAVG
 51  ALASVVFLLL LARGFPRVSS VVLLIYVGTT ALYLPVGWLY GAPSYQIVGS
101  ILESNPAEAR EFVGNLPGSL YFVQALFFIF GLTVWKYCVS GGVFADVKNY
151  KRRSKIWLTI LLTLILSCAV MDKIASDKDL REPDAGLLLN IFDLYYDLAX
201  RAGTICRQAR PHFGSSKKSV NMAYPSCCAQ V*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 084 shows 90.5% identity over a 231 aa overlap with a predicted ORF (ORF 084.ng) from *N. gonorrhoeae:*

```
m084/g084

                    10        20                 30        40        50
m084.pep    MKQSARIKXMNQTLLYTLGICALLTF---------YHPEYEYGYRYSAVGALASVVFLLL
            |||||||| |:|||  ||||||||:|         || ||||||||||||||||||||||
g084        MKQSARIKNMDQTLKNTLGICALLAFCFGAAIASGYHLEYEYGYRYSAVGALASVVFLLL
                    10        20        30        40        50        60

                   60        70        80        90       100       110
m084.pep    LARGFPRVSSVVLLIYVGTTALYLPVGWLYGAPSYQIVGSILESNPAEAREFVGNLPGSL
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g084        LARGFPRVSSVVLLIYVGTTALYLPVGWLYGAPSYQIVGSILESNPAEAREFVGNLPGSL
                    70        80        90       100       110       120

                  120       130       140       150       160       170
m084.pep    YFVQALFFIFGLTVWKYCVSGGVFADVKNYKRRSKIWLTILLTLILSCAVMDKIASDKDL
            |||||||||||||||||||| ||||||||||||||||||||||||||||||:|||:||| 
g084        YFVQALFFIFGLTVWKYCVSVGVFADVKNYKRRSKIWLTILLTLILSCAVMEKIAGDKDW
                   130       140       150       160       170       180
```

```
                  180       190       200       210       220
m084.pep  REPDAGLLLNIFDLYYDLAXRAGTICRQARPHFGSSKKSVNMAYPSCCAQVX
          ||||||||||||||||||| |||||||||||||||||||||||||  |||||
g084      REPDAGLLLNIFDLYYDLAFRAGTICRQARPHFGSSKKSVNMAYPPTCAQVX
                   190       200       210       220       230
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 263>:

a084.seq

```
  1  ATGAAACAAT CCGCCCGAAT AAAAAATATG GATCAGACAT TAAAAAATAC
 51  ATTGGGCATT TGCGCGCTTT TAGCCTTTTG TTTTGGCGCG GCCATCGCAT
101  CAGGTTATCA CTTGGAATAT GAATACGGCT ACCGTTATTC TGCCGTGGGT
151  GCTTTGGCTT CGGTTGTATT TTTATTATTA TTGGCACGCG GTTTCCCGCG
201  CGTTTCTTCA GTTGTTTTAC TGATTTACGT CGGCACAACC GCCCTATATT
251  TGCCGGTCGG CTGGCTGTAT GGTGCGCCGT CTTATCAGAT AGTCGGTTCG
301  ATATTGGAAA GCAATCCTGC CGAGGCGCGT GAATTTGTCG GCAATCTTCC
351  CGGGTCGCTT TATTTTGTGC AGGCATTATT TTTCATTTTT GGCTTGACAG
401  TTTGGAGATA TTGTGTATCG GGGGGGGTAT TTGCTGACGT AAAAAACTAT
451  AAACGCCGCA GCAAAAATATG GCTGACTATA TTATTGACTT TGATTTTGTC
501  CTGCGCGGTG ATGGATAAAA TCGCCAGCGA TAAAGATTTG CGAGAACCTG
551  ATGCCGGCCT GTTGTTGAAT ATTTTCGACC TGTATTACGA TTTGGCTTCC
601  .GCGCCGGCA CAATATGCCG CCAAGCGCGC CCACATTTTG GAAGCAGCAA
651  AAAAAGCGTC AACATGGCAT ATCCGTCATG TTGCGCCCAA GTATAA
```

This corresponds to the amino acid sequence <SEQ ID 264; ORF 084.a>:

a084.pep

```
  1  MKQSARIKNM DQTLKNTLGI CALLAFCFGA AIASGYHLEY EYGYRYSAVG
 51  ALASVVFLLL LARGFPRVSS VVLLIYVGTT ALYLPVGWLY GAPSYQIVGS
101  ILESNPAEAR EFVGNLPGSL YFVQALFFIF GLTVWRYCVS GGVFADVKNY
151  KRRSKIWLTI LLTLILSCAV MDKIASDKDL REPDAGLLLN IFDLYYDLAS
201  XAGTICRQAR PHFGSSKKSV NMAYPSCCAQ V*
```

m084/a084 92.2% identity over a 231 aa overlap

```
                  10        20        30        40        50        60
m084.pep  MKQSARIKXMNQTLLYTLGICALLTFXXXXXXXXXYHPEYEYGYRYSAVGALASVVFLLL
          |||||||| |:|||  ||||||||:|          || |||||||||||||||||||||
a084      MKQSARIKNMDQTLKNTLGICALLAFCFGAAIASGYHLEYEYGYRYSAVGALASVVFLLL
                  10        20        30        40        50        60

                  70        80        90       100       110       120
m084.pep  LARGFPRVSSVVLLIYVGTTALYLPVGWLYGAPSYQIVGSILESNPAEAREFVGNLPGSL
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a084      LARGFPRVSSVVLLIYVGTTALYLPVGWLYGAPSYQIVGSILESNPAEAREFVGNLPGSL
                  70        80        90       100       110       120

                 130       140       150       160       170       180
m084.pep  YFVQALFFIFGLTVWKYCVSGGVFADVKNYKRRSKIWLTILLTLILSCAVMDKIASDKDL
          |||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
a084      YFVQALFFIFGLTVWRYCVSGGVFADVKNYKRRSKIWLTILLTLILSCAVMDKIASDKDL
                 130       140       150       160       170       180

                 190       200       210       220       230
m084.pep  REPDAGLLLNIFDLYYDLAXRAGTICRQARPHFGSSKKSVNMAYPSCCAQVX
          |||||||||||||||||||  |||||||||||||||||||||||||||||||
a084      REPDAGLLLNIFDLYYDLASXAGTICRQARPHFGSSKKSVNMAYPSCCAQVX
                 190       200       210       220       230
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 265>:

g085.seq

```
  1  ATGGGCAAAG GGCAGGACTT CACGCCCCTG CGCGACGCGT TGAAAGATAA
```

```
 51  GGCAAAAGGC GTGTTCCTGA TCGGCGTCGA TGCGCCGCAA ATCCGCCGCG
101  ATTTGGACGG CTGCGGCTTG AACCTGACCG ACTGCGTCAC TTTGGAAGAG
151  GCGGTTCAGA CGGCATACGC CCAAGCCGAA GCGGGCGATA TTGTCTTGCT
201  CAGCCCCGCC TGCGCGAGTT TCGATATGTT TAAAGGCTAC GCGCACCGTT
251  CGGAAGTGTT tatCGAAGCG TTTAAGGCTT TGTGA
```

This corresponds to the amino acid sequence <SEQ ID 266; ORF 085.ng>:

g085.pep

```
  1  MGKGQDFTPL RDALKDKAKG VFLIGVDAPQ IRRDLDGCGL NLTDCVTLEE
 51  AVQTAYAQAE AGDIVLLSPA CASFDMFKGY AHRSEVFIEA FKAL*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 267>:

m085.seq

```
  1  ATGGGTAAAG GGCAGGACTT CACGCCCCTG CGCGATGCAC TGGTAGGCAA
 51  GGCAAAAGGC GTGTTCTTGA TTGGTGTCGA TGCGCCGCAA ATCCGCCGCG
101  ATTTGGACGG CTGCGGCTTG AATATGACCG ACTGCGCCAC TTTGGGAGAA
151  GCCGTTCAGA CGGCATATGC CCAAGCCGAA GCAGGCGATA TTGTGTTGCT
201  CAGCCCCGCC TGCGCGAGCT TTGATATGTT CAAAGGCTAC GCGCACCGTT
251  CGGAAGTGTT TATCGAAGCG TTTAAGGCTT TGTGA
```

This corresponds to the amino acid sequence <SEQ ID 268; ORF 085>:

m085.pep

```
  1  MGKGQDFTPL RDALVGKAKG VFLIGVDAPQ IRRDLDGCGL NMTDCATLGE
 51  AVQTAYAQAE AGDIVLLSPA CASFDMFKGY AHRSEVFIEA FKAL*
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

ORF 085 shows 94.7% identity over a 94 aa overlap with a predicted ORF (ORF 085.ng) from *N. gonorrhoeae:*

m085/g085

```
                     10        20        30        40        50        60
m085.pep    MGKGQDFTPLRDALVGKAKGVFLIGVDAPQIRRDLDGCGLNMTDCATLGEAVQTAYAQAE
            ||||||||||||||  |||||||||||||||||||||||||:|||:|| |||||||||||
g085        MGKGQDFTPLRDALKDKAKGVFLIGVDAPQIRRDLDGCGLNLTDCVTLEEAVQTAYAQAE
                     10        20        30        40        50        60

                     70        80        90
m085.pep    AGDIVLLSPACASFDMFKGYAHRSEVFIEAFKALX
            |||||||||||||||||||||||||||||||||||
g085        AGDIVLLSPACASFDMFKGYAHRSEVFIEAFKALX
                     70        80        90
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 269>:

a085.seq

```
  1  ATGGGCAAAG GGCAGGACTT CACGCCCCTG CGCGACGCGC TTGCCGGCAA
 51  GGCAAAAGGC GTGTTCCTGA TCGGTGTCGA TGCGCCGCAA ATCCGCCGCG
101  ATTTGGACGG CTGCGATCTG AATATGACCG ACTGCGCCAC TTTGGAAGAA
151  GCGGTTCAGA AGGCATATGC CCAAGCCGAA GCGGGCGATA TCGTGCTGCT
201  CAGCCCCGCC TGCGCGAGTT TCGATATGTT TAAAGGCTAC GCGCACCGTT
251  CGGAAGTGTT TATCGGGGCG TTTAAGGCTT TGTGA
```

This corresponds to the amino acid sequence <SEQ ID 270; ORF 085.a>:

a085.pep

```
  1  MGKGQDFTPL RDALAGKAKG VFLIGVDAPQ IRRDLDGCDL NMTDCATLEE
 51  AVQKAYAQAE AGDIVLLSPA CASFDMFKGY AHRSEVFIGA FKAL*
```

m085/a085 94.7% identity over a 94 aa overlap

```
                     10        20        30        40        50        60
m085.pep    MGKGQDFTPLRDALVGKAKGVFLIGVDAPQIRRDLDGCGLNMTDCATLGEAVQTAYAQAE
```

```
              ||||||||||||||:|||||'''||||||||||||||| ||||||||| |||| ||||||
a085          MGKGQDFTPLRDALAGKAKGVFLIGVDAPQIRRDLDGCDLNMTDCATLEEAVQKAYAQAE
                        10        20        30        40        50        60

                        70        80        90
m085.pep      AGDIVLLSPACASFDMFKGYAHRSEVFIEAFKALX
              |||||||||||||||||||||||||||| ||.|||
a085          AGDIVLLSPACASFDMFKGYAHRSEVFIGAFKALX
                        70        80        90
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 271>:

```
g086.seq
   1  ATGGTGGTGC TGATGACGGC GTTCGGCCTG CTGATGATTT ATTCGGCTTC
  51  TGTGTATTTG GCATCGAAGG AAGGCGGCGA TCAGTTTTTC TATTTGACCA
 101  GGCAGGCGGG GTTCGTCGTT GCCGGCCTTA TAGCGAGCGG TTTTTATGG
 151  TTTCTTTGCA GGATGAGGAC ATGGCGGCGG CTTGTGCCGT GGATTTTTGC
 201  CTTATCCGGC CTGTTGCTGG TAGCCGTATT GATTGCCGGG CGCGAAATCA
 251  ATGGCGCGAC CCGTTGGATA CCTTTGGGTC CGTTGAATTT CCAGCCGACC
 301  GAGCTGTTCA AGCTGGCAGT CATCCTTTAT TTGGCAAGCC TGTTCACGCG
 351  CCGTGAAGAA GTGTTGCGCA GCATGGAAAG TTTGGGTTGG CAGTCGATTT
 401  GGCGGGGGAC GGCCAACCTG ATTATGTCCG CCACCAATCC GCAGGCACGT
 451  CGTGAAACAT TAGAAATGTA CGgcCGTTTC CGGGCGATCA TCCTGCCGAT
 501  TATGCTGGTG GCGTTCGGTT TGGTGCTGAT AATGGTACAG CCGGATTTCG
 551  GTTCGTTTGT CGTCATTACC GTCATTACCG TTGGAATGCT GTTTCTGGCA
 601  GGATTGCCGT GGAAATATTT TTTTGTCCTG GTAGGCAGCG TCTTGGGTGG
 651  GATGGTGCTG ATGATTACCG CCGCTCCCTA CCGTGTGCAG CGGGTAGTGG
 701  CATTTTTGGA CCCGTGGAAA GACCCGCAGG GTGCCGGCTA CCAGCTTACC
 751  CACTCTCTGA TGGCAATCGG GCGCGGAGAG TGGTTCGGTA TGGGTTTGGG
 801  TGCGAGTTTG AGCAAACGCG GCTTTCTGCC GGAAGCGCAT ACCGATTTTA
 851  TTTTTGCCAT CATCGCTGAA GAATTCGGCT TCTTCGGGAT GTGCGTGCTG
 901  ATATTCTGTT ACGGCTGGCT GGTGGTGCGG GCGTTTTCCA TCGGCAAGCA
 951  GTCGCGCGAT TTGGGtttgA CTTTCAACGC CTATATCGCT TCGGGTATCG
1001  GCATTTGGAT CGGTATCCAA AGTTTCTTCA ATATCGGTGT GAACATCGGT
1051  GCTTTGCCGA CCAAAGGTCT GACGctgCcg tTGATGTCCT ATGGcggTTC
1101  GTCAGTCTTT TTCATGCTGA TCAGCATGAT GCTGCTGTTG CGTATCGATT
1151  ATGAAAACCG CCAGAAAATG CGCGGTTACC GGGTGGAGTA AA
```

This corresponds to the amino acid sequence <SEQ ID 272; ORF 086.ng>:

```
g086.pep
   1  MVVLMTAFGL LMIYSASVYL ASKEGGDQFF YLTRQAGFVV AGLIASGFLW
  51  FLCRMRTWRR LVPWIFALSG LLLVAVLIAG REINGATRWI PLGPLNFQPT
 101  ELFKLAVILY LASLFTRREE VLRSMESLGW QSIWRGTANL IMSATNPQAR
 151  RETLEMYGRF RAIILPIMLV AFGLVLIMVQ PDFGSFVVIT VITVGMLFLA
 201  GLPWKYFFVL VGSVLGGMVL MITAAPYRVQ RVVAFLDPWK DPQGAGYQLT
 251  HSLMAIGRGE WFGMGLGASL SKRGFLPEAH TDFIFAIIAE EFGFFGMCVL
 301  IFCYGWLVVR AFSIGKQSRD LGLTFNAYIA SGIGIWIGIQ SFFNIGVNIG
 351  ALPTKGLTLP LMSYGGSSVF FMLISMMLLL RIDYENRQKM RGYRVE*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 273>:

```
m086.seq
   1  ATGGTGGTGC TGATGACGGC GTTCAGCCTG CTGATGATTT ATTCGGCTTC
  51  TGTGTATTTG GCATCAAAAG AAGGCGGCGA TCAGTTTTTC TATTTGACCA
 101  GACAGGCGGG GTTCGTCGTT GCCGGCTTGA TAGCGAGCGG TTTGTTATGG
 151  TTTCTTTGCA GGATGAGGAC ATGGCGGCGG CTTGTGCCGT GGATTTTTGC
 201  CCTATCCGGC CTGTTGCTGG TAGTCGTATT GATTGCCGGG CGCGAAATCA
 251  ATGGCGCGAC CCGTTGGATA CCTTTGGGTC CGTTGAATTT CCAGCCGACc
 301  GAGCTGTTCA AGCtGGCGGT CATCCTTTAT TTGGCAAGCC TGTTCACGCG
 351  CCGTGAAGAA GTGTTGcGCA GCATGGAAAG TTTGGGTTGG CAGTCGATTT
 401  GGCGGGGGAC GGCCAATCTG ATCATGTCCG CCACCAATCC GCAGrCACGT
 451  CGTGAaACAT TAGAAATGTA CGGCCGTwTC CGGGCGATCA TCCTGCCGAT
 501  TATGCTGGTG GCGTTCGGTT TGGTGCTGAT AATGGTACAG CCGGATTTCG
```

```
551   GTTCGTTTGT CGTCATTACC GTCATTGCCG TTGGAATGCT GTTTTTGGCA
601   GGATTGCCGT GGAAATATTT TTTCGTCCTG GTAGGCAGCG TCTTGGGCGG
651   GATGGTGCTG ATGATTACCG CCGCTCCCTA CCGTGTGCAG CGGGTAGTGG
701   CATTTTTGGA CCCGTGGAAA GACCCGCAGG GTGCCGGCTA CCAGCTTACC
751   CACTCTCTGA TGGCAATCGG GCGCGGAGAG TGGTTCGGTA TGGGTTTGGG
801   TGCGAGTTTG AGCAAACGCG GCTTTCTGCC GGAAGCGCAT ACCGATTTTA
851   TTTTTGCCAT CATCGCCGAA GAATTCGGTT TCTTCGGTAT GTGCGTGCTG
901   ATATTCTGTT ACGGCTGGCT GGTGGTGCGG GCGTTTTCCA TCGGCAAGCA
951   GTCGCGCGAT TTGGGTTTGA CTTTCAACGC CTATATCGCT TCGGGTATCG
1001  GCATTTGGAT CGGkrTCCAA AGTTTCTTCA ATATCGGTGT GAACATCGGT
1051  GCTTTGCCGA mCAAAgGyCT GACGCyGCCG Tg.AtGTCCw ATGGCGGTTC
1101  GTCAGTCTTT TTCATGCTGA TCAGCATGAT GCTGCTGTkG CGTATAGATT
1151  ATGAAAACCG CCGGAAAATG CGCGGTTATC GGGTGGAGTA A
```

This corresponds to the amino acid sequence <SEQ ID 274; ORF 086>:

```
m086.pep
  1  MVVLMTAFSL LMIYSASVYL ASKEGGDQFF YLTRQAGFVV AGLIASGLLW
 51  FLCRMRTWRR LVPWIFALSG LLLVVVLIAG REINGATRWI PLGPLNFQPT
101  ELFKLAVILY LASLFTRREE VLRSMESLGW QSIWRGTANL IMSATNPQXR
151  RETLEMYGRX RAIILPIMLV AFGLVLIMVQ PDFGSFVVIT VIAVGMLFLA
201  GLPWKYFFVL VGSVLGGMVL MITAAPYRVQ RVVAFLDPWK DPQGAGYQLT
251  HSLMAIGRGE WFGMGLGASL SKRGFLPEAH TDFIFAIIAE EFGFFGMCVL
301  IFCYGWLVVR AFSIGKQSRD LGLTFNAYIA SGIGIWIGXQ SFFNIGVNIG
351  ALPXKGLTXP XMSXGGSSVF FMLISMMLLX RIDYENRRKM RGYRVE*
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

ORF 086 shows 96.7% identity over a 396 aa overlap with a predicted ORF (ORF 086.ng) from *N. gonorrhoeae:*

```
m086/g086

                   10        20        30        40        50        60
m086.pep   MVVLMTAFSLLMIYSASVYLASKEGGDQFFYLTRQAGFVVAGLIASGLLWFLCRMRTWRR
           ||||||||:||||||||||||||||||||||||||||||||||||||:||||||||||||
g086       MVVLMTAFGLLMIYSASVYLASKEGGDQFFYLTRQAGFVVAGLIASGFLWFLCRMRTWRR
                   10        20        30        40        50        60

                   70        80        90       100       110       120
m086.pep   LVPWIFALSGLLLVVVLIAGREINGATRWIPLGPLNFQPTELFKLAVILYLASLFTRREE
           ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
g086       LVPWIFALSGLLLVAVLIAGREINGATRWIPLGPLNFQPTELFKLAVILYLASLFTRREE
                   70        80        90       100       110       120

                  130       140       150       160       170       180
m086.pep   VLRSMESLGWQSIWRGTANLIMSATNPQXRRETLEMYGRXRAIILPIMLVAFGLVLIMVQ
           |||||||||||||||||||||||||||| ||||||||||| ||||||||||||||||||||
g086       VLRSMESLGWQSIWRGTANLIMSATNPQARRETLEMYGRFRAIILPIMLVAFGLVLIMVQ
                  130       140       150       160       170       180

                  190       200       210       220       230       240
m086.pep   PDFGSFVVITVIAVGMLFLAGLPWKYFFVLVGSVLGGMVLMITAAPYRVQRVVAFLDPWK
           ||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
g086       PDFGSFVVITVITVGMLFLAGLPWKYFFVLVGSVLGGMVLMITAAPYRVQRVVAFLDPWK
                  190       200       210       220       230       240

                  250       260       270       280       290       300
m086.pep   DPQGAGYQLTHSLMAIGRGEWFGMGLGASLSKRGFLPEAHTDFIFAIIAEEFGFFGMCVL
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g086       DPQGAGYQLTHSLMAIGRGEWFGMGLGASLSKRGFLPEAHTDFIFAIIAEEFGFFGMCVL
                  250       260       270       280       290       300

                  310       320       330       340       350       360
```

```
m086.pep    IFCYGWLVVRAFSIGKQSRDLGLTFNAYIASGIGIWIGXQSFFNIGVNIGALPXKGLTXP
            |||||||||||||||||||||||||||||||||||||||| |||||||||||||:|||| |
g086        IFCYGWLVVRAFSIGKQSRDLGLTFNAYIASGIGIWIGIQSFFNIGVNIGALPTKGLTLP
                    310       320       330       340       350       360

                    370       380       390
m086.pep    XMSXGGSSVFFMLISMMLLXRIDYENRRKMRGYRVEX
             || |||||||||||||| ||||||:||||||||
g086        LMSYGGSSVFFMLISMMLLLRIDYENRQKMRGYRVEX
                    370       380       390
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 275>:

a086.seq

```
   1  ATGGTGGTGC TGATGACGGC GTTCAGCCTG CTGATGATTT ATTCGGCTTC
  51  TGTGTATTTG GCATCAAAAG AAGGCGGCGA TCAGTTTTTC TATTTGACCA
 101  GACAGGCGGG GTTCGTCGTT GCCGGCTTGA TAGCGAGCGG TTTGTTATGG
 151  TTTCTTTGCA GGATGAGGAC ATGGCGGCGG CTTGTGCCGT GGATTTTTGC
 201  CCTATCCGGC CTGTTGCTGG TAGTCGTATT GATTGCCGGG CGCGAAATCA
 251  ATGGCGCGAC CCGTTGGATA CCTTTGGGTC CGTTGAATTT CCAGCCGACC
 301  GAGCTGTTCA AGCTGGCGGT CATCCTTTAT TTGGCAAGCC TGTTCACGCG
 351  CCGTGAAGAA GTGTTGCGCA GCATGGAAAG TTTGGGTTGG CAGTCGATTT
 401  GGCGGGGGAC GGCCAATCTG ATCATGTCCG CCACCAATCC GCAGGCACGT
 451  CGTGAAACAT TAGAAATGTA CGGCCGTTTC CGGGCGATCA TCCTGCCGAT
 501  TATGCTGGTG GCGTTCGGTT TGGTGCTGAT AATGGTACAG CCGGATTTCG
 551  GTTCCTTTGT CGTCATTACC GTCATTGCCG TTGGAATGCT GTTTTTGGCA
 601  GGATTGCCGT GGAAATATTT TTTCGTCCTG GTAGGCAGCG TCTTGGGCGG
 651  GATGGTGCTG ATGATTACCG CCGCTCCCTA CCGTGTGCAG CGGGTAGTGG
 701  CATTTTTGGA CCCGTGGAAA GACCCGCAGG GTGCCGGCTA CCAGCTTACC
 751  CACTCTCTGA TGGCAATCGG GCGCGGAGAG TGGTTCGGTA TGGGTTTGGG
 801  TGCGAGTTTG AGCAAACGCG GCTTTCTGCC GGAAGCGCAT ACCGATTTTA
 851  TTTTTGCCAT CATCGCCGAA GAATTCGGTT TCTTCGGTAT GTGCGTGCTG
 901  ATATTCTGTT ACGGCTGGCT GGTGGTGCGG GCGTTTTCCA TCGGCAAGCA
 951  GTCGCGCGAT TTGGGTTTGA CTTTCAACGC CTATATCGCT TCGGGTATCG
1001  GCATTTGGAT CGGTATCCAA AGTTTCTTCA ATATCGGTGT GAACATCGGT
1051  GCTTTGCCGA CCAAAGGTCT GACGCTGCCG TTGATGTCCT ATGGCGGTTC
1101  GTCAGTCTTT TTCATGCTGA TCAGCATGAT GCTGCTGTTG CGTATAGATT
1151  ATGAAAACCG CCGGAAAATG CGCGGTTACC GGGTGGAGTA A
```

This corresponds to the amino acid sequence <SEQ ID 276; ORF 086.a>:

a086.pep

```
   1  MVVLMTAFSL LMIYSASVYL ASKEGGDQFF YLTRQAGFVV AGLIASGLLW
  51  FLCRMRTWRR LVPWIFALSG LLLVVVLIAG REINGATRWI PLGPLNFQPT
 101  ELFKLAVILY LASLFTRREE VLRSMESLGW QSIWRGTANL IMSATNPQAR
 151  RETLEMYGRF RAIILPIMLV AFGLVLIMVQ PDFGSFVVIT VIAVGMLFLA
 201  GLPWKYFFVL VGSVLGGMVL MITAAPYRVQ RVVAFLDPWK DPQGAGYQLT
 251  HSLMAIGRGE WFGMGLGASL SKRGFLPEAH TDFIFAIIAE EFGFFGMCVL
 301  IFCYGWLVVR AFSIGKQSRD LGLTFNAYIA SGIGIWIGIQ SFFNIGVNIG
 351  ALPTKGLTLP LMSYGGSSVF FMLISMMLLL RIDYENRRKM RGYRVE*
```

m086/a086 98.0% identity over a 396 aa overlap

```
                    10        20        30        40        50        60
m086.pep    MVVLMTAFSLLMIYSASVYLASKEGGDQFFYLTRQAGFVVAGLIASGLLWFLCRMRTWRR
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a086        MVVLMTAFSLLMIYSASVYLASKEGGDQFFYLTRQAGFVVAGLIASGLLWFLCRMRTWRR
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m086.pep    LVPWIFALSGLLLVVVLIAGREINGATRWIPLGPLNFQPTELFKLAVILYLASLFTRREE
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a086        LVPWIFALSGLLLVVVLIAGREINGATRWIPLGPLNFQPTELFKLAVILYLASLFTRREE
```

```
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m086.pep    VLRSMESLGWQSIWRGTANLIMSATNPQXRRETLEMYGRXRAIILPIMLVAFGLVLIMVQ
            |||||||||||||||||||||||||||| ||||||||| |||||||||||||||||||||
a086        VLRSMESLGWQSIWRGTANLIMSATNPQARRETLEMYGRFRAIILPIMLVAFGLVLIMVQ
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m086.pep    PDFGSFVVITVIAVGMLFLAGLPWKYFFVLVGSVLGGMVLMITAAPYRVQRVVAFLDPWK
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a086        PDFGSFVVITVIAVGMLFLAGLPWKYFFVLVGSVLGGMVLMITAAPYRVQRVVAFLDPWK
                   190       200       210       220       230       240

                   250       260       270       280       290       300
m086.pep    DPQGAGYQLTHSLMAIGRGEWFGMGLGASLSKRGFLPEAHTDFIFAIIAEEFGFFGMCVL
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a086        DPQGAGYQLTHSLMAIGRGEWFGMGLGASLSKRGFLPEAHTDFIFAIIAEEFGFFGMCVL
                   250       260       270       280       290       300

                   310       320       330       340       350       360
m086.pep    IFCYGWLVVRAFSIGKQSRDLGLTFNAYIASGIGIWIGXQSFFNIGVNIGALPXKGLTXP
            ||||||||||||||||||||||||||||||||||||||| |||||||||||||:|||| |
a086        IFCYGWLVVRAFSIGKQSRDLGLTFNAYIASGIGIWIGIQSFFNIGVNIGALPTKGLTLP
                   310       320       330       340       350       360

                   370       380       390
m086.pep    XMSXGGSSVFFMLISMMLLXRIDYENRRKMRGYRVEX
             || ||||||||||||||| |||||||||||||||||
a086        LMSYGGSSVFFMLISMMLLLRIDYENRRKMRGYRVEX
                   370       380       390
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 277>:

```
g087.seq
       1  ATGGGCGGTA AAACCTTTAT GCTGATGGCG GGCGGAACGG GCGGACACAT
      51  TTTCCCAGCT CTGGCTGTGG CGGATTCATT GCGCGTGCGC GGTCATCATG
     101  TAATTTGGCT GGGCAGCAAG GATTCGATGG AAGAGCGCAT CGTGCCGCAA
     151  TACGGCATAC GCTTGGAAAC GCTGGCGATT AAAGGAATAC GCGGCAACGG
     201  CATCAAACGC AAGCTGATGC TTCCGTTTAC TCTGTACAAA ACCGTCCGCG
     251  AAGCGCAGCG GATTATCCGC AAACACCGTG TCGAGTGCGT CATCGGCTTC
     301  GGCGGTTTTG TTACCTTTCC CGGCGGTCTG GCGGCGAAAC TCTTGGGCGT
     351  GCCGATTGTG ATTCACGAGC AAAACGCCGT GGCAGGCTTG TCCAACCGCC
     401  AccTGTCGCg ctGGGCGAAA CGGGTGTTGT ACGCTTTTCC GAAAGCGTTC
     451  AGCCACGAAG GCGGTTTGGT CGGCAACCCC GTCCGCGCCG ATATTAGCAA
     501  CCTGCCCGTG CCTGCCGAAC GCTTCCAAGG GCGCGAAGGC CGTCTGAAAA
     551  TTTTGGTGGT CGGCGGCAGT TTGGGTGCGG ACGTTTTGAA CAAAACCGTA
     601  CCGCAGGCGT TGGCACTGCT GCCTGAAGAG GTGCGCCCGC AGATGTACCA
     651  CCAGTCGGGG CGTAACAAGC TGGGCAATCT TCAGGCGGAT TATGACGCGT
     701  TGGGCGTGAA AGCGGAATGC GTGGAATTTA TTACCGACAT GGTGTCCGCC
     751  TACCGTGATG CCGATTTGGT GATTTGCCGT GCCGGCGCGC TGACGATTGC
     801  CGAGTTGACG GCGGCGGGGC TGGGCGCGTT GTTAGTGCCG TATCCTCACG
     851  CCGTTGATGA CCATCAAACC GCCAACGCGC GTTTCATGGT GCAGGCAGAA
     901  GCGGGGCTGC TGTTGCCGCA AACCCAGTTG ACGGCGGAAA AACTCGCCGA
     951  AATCCTCGGC AGCCTCAACC GCGAAAAATG CCTCAAATGG GCGGAAAACG
    1001  CCCGTACGTT GGCATTGCCG CACAGCGCGG ATGACGTTGC CGAAGCCGCG
    1051  ATTGCGTGTG CGGCGTAAA
```

This corresponds to the amino acid sequence <SEQ ID 278; ORF 087.ng>:

```
g087.pep
       1  MGGKTFMLMA GGTGGHIFPA LAV[illegible]SLRVR GHHVIWLGSK DSMEERIVPQ
      51  YGIRLETLAI KGIRGNGIKR KLMLPFTLYK TVREAQRIIR KHRVECVIGF
     101  GGFVTFPGGL AAKLLGVPIV IHEQNAVAGL SNRHLSRWAK RVLYAFPKAF
     151  SHEGGLVGNP VRADISNLPV PAERFQGREG RLKILVVGGS LGADVLNKTV
```

```
201  PQALALLPEE VRPQMYHQSG RNKLGNLQAD YDALGVKAEC VEFITDMVSA
251  YRDADLVICR AGALTIAELT AAGLGALLVP YPHAVDDHQT ANARFMVQAE
301  AGLLLPQTQL TAEKLAEILG SLNREKCLKW AENARTLALP HSADDVAEAA
351  IACAA*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 279>:

```
m087.seq
   1  ATGGGCGGTA AAACCTTTAT GCTGAwkkCG GGCGGAACGG GCGGACATAT
  51  TTTCCCCGCG CTGGCGGTGG CGGATTCATT GCGCGCGCGC GGCCATCATG
 101  TGATTTGGCT GGGCAGCAAG GATTCGATGG AAGAGCGTAT CGTGCCGCAA
 151  TACGGCATAC GCTTGGAAAC GCTGGCGATT AAAGGCGTGC GCGGCAACGG
 201  CATCAAACGC AAACTGATGC TGCCGGTTAC TTTGTATCAA ACCGTCCGCG
 251  AAGCGCAGCG GATTATCCGC AAACACCGTG TCGAGTGCGT CATCGCCTTC
 301  GGCGGCTTCG TTACCTTCCC CGGCGGTTTG GCGGCGAAGC TATTArGCGT
 351  GCCGATTGTG ATTCACGAGC AAAACGCCGT GGCAGGTTTG TCCAACCGCC
 401  ACCTGTCGCG CTGGGCGAAG CGGGTGTTGT ACGCTTTTCC GAAAGCGTTC
 451  AGCCACGAAG GCGGCTTGGT CGGCAACCCC GTCCGCGCCG ATATTAGCAA
 501  CCTGCCCGTG CCTGCCGAAC GCTTCCAAGG GCGTGAAGGC CGTCTGAAAA
 551  TTTTGGTGGT CGGCGGCAGT TTGGGCGCGG ACGTTTTGAA CAAAACCGTA
 601  CCGCATGCAT TGGCTTTGCT GCCCGACAAT GCGCGTCCGC ATATGTACCA
 651  CCAATCGGGA CGGGGCAAGC TGGGCATCTT GCAGGCGnnn nnnnnnnnnn
 701  nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 751  nnnGCGGGAT TGGGTGCGTT GTTAGTGCCG TATCCTCACG CGGTTGACGA
 801  TCACCAAACC GCCAACGCGC GTTTTATGGT GCAGGCGGAG GCGGGATTGC
 851  TGTTGCCGCA AACCCAGTTG ACGGCGGAAA AACTCGCCGA GATTCTCGGC
 901  GGCTTAAACC GCGAAAAATG CCTCAAATGG GCAGAAAACG CCCGTACGTT
 951  GGCACTGCCG CACAGTGCGG ACGACGTGGC GGAAGCCGCG ATTGCGTGTG
1001  CGGCGTAA
```

This corresponds to the amino acid sequence <SEQ ID 280; ORF 087>:

```
m087.pep
   1  MGGKTFMLXX GGTGGHIFPA LAVADSLRAR GHHVIWLGSK DSMEERIVPQ
  51  YGIRLETLAI KGVRGNGIKR KLMLPVTLYQ TVREAQRIIR KHRVECVIGF
 101  GGFVTFPGGL AAKLLXVPIV IHEQNAVAGL SNRHLSRWAK RVLYAFPKAF
 151  SHEGGLVGNP VRADISNLPV PAERFQGREG RLKILVVGGS LGADVLNKTV
 201  PHALALLPDN ARPHMYHQSG RGKLGILQAX XXXXXXXXXX XXXXXXXXXX
 251  XAGLGALLVP YPHAVDDHQT ANARFMVQAE AGLLLPQTQL TAEKLAEILG
 301  GLNREKCLKW AENARTLALP HSADDVAEAA IACAA*
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

ORF 087 shows 83.9% identity over a 355 aa overlap with a predicted ORF (ORF 087.ng) from *N. gonorrhoeae:*

```
m087/g087

                    10        20        30        40        50        60
m087.pep    MGGKTFMLXXGGTGGHIFPALAVADSLRARGHHVIWLGSKDSMEERIVPQYGIRLETLAI
            ||||||||  |||||||||||||||||||:||||||||||||||||||||||||||||||
g087        MGGKTFMLMAGGTGGHIFPALAVADSLRVRGHHVIWLGSKDSMEERIVPQYGIRLETLAI
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m087.pep    KGVRGNGIKRKLMLPVTLYQTVREAQRIIRKHRVECVIGFGGFVTFPGGLAAKLLXVPIV
            ||:||||||||||||| |||:||||||||||||||||||||||||||||||||||| ||||
g087        KGIRGNGIKRKLMLPFTLYKTVREAQRIIRKHRVECVIGFGGFVTFPGGLAAKLLGVPIV
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m087.pep    IHEQNAVAGLSNRHLSRWAKRVLYAFPKAFSHEGGLVGNPVRADISNLPVPAERFQGREG
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g087        IHEQNAVAGLSNRHLSRWAKRVLYAFPKAFSHEGGLVGNPVRADISNLPVPAERFQGREG
                   130       140       150       160       170       180
```

```
                 190       200       210       220       229
m087.pep   RLKILVVGGSLGADVLNKTVPHALALLPDNARPHMYHQSGRGKLGILQA----------
           ||||||||||||||||||||:||||||:::||:|||||||:||| |||
g087       RLKILVVGGSLGADVLNKTVPQALALLPEEVRPQMYHQSGRNKLGNLQADYDALGVKAEC
                 190       200       210       220       230       240

                                             230       240       250
m087.pep   -----------------------------AGLGALLVPYPHAVDDHQTANARFMVQAE
                                         |||||||||||||||||||||||||||||
g087       VEFITDMVSAYRDADLVICRAGALTIAELTAAGLGALLVPYPHAVDDHQTANARFMVQAE
                 250       260       270       280       290       300

          260       270       280       290       300       310
m087.pep   AGLLLPQTQLTAEKLAEILGGLNREKCLKWAENARTLALPHSADDVAEAAIACAAX
           |||||||||||||||||||:||||||||||||||||||||||||||||||||||||
g087       AGLLLPQTQLTAEKLAEILGSLNREKCLKWAENARTLALPHSADDVAEAAIACAAX
                 310       320       330       340       350
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 281>:

a087.seq

```
   1  ATGGGCGGTA AAACCTTTAT GCTGATGGCG GGCGGAACGG GCGGACATAT
  51  TTTCCCCGCG CTGGCGGTGG CGGATTCATT GCGCGCGCGC GGCCATCATG
 101  TAATTTGGCT GGGCAGCAAG GATTCGATGG AAGAGCGCAT CGTGCCGCAA
 151  TACGACATCC TGCTCGAAAC GCTGGCGATT AAAGGCGTGC GCGGCAACGG
 201  CATCAAACGC AAGCTGATGC TGCCGTTTAC TTTGTATCAA ACTGTCCGCG
 251  AAGCGCAGCA GATTATCCGC AAACACCGTG TCGAGTGCGT CATCGGCTTC
 301  GGCGGCTTCG TTACCTTTCC CGGCGGTTTG GCGGCGAAGT TATTAGGCGT
 351  GCCGATTGTG ATTCACGAGC AAAACGCCGT GGCAGGTTTG TCCAACCGCC
 401  ACCTGTCGCG CTGGGCGAAG CGGGTGTTGT ACGCTTTTCC GAAAGCGTTC
 451  AGCCACGAAG GCGGCTTGGT CGGCAACCCC GTCCGCGCCG ATATTAGCAA
 501  CCTGCCCGTG CCTGCCGAAC GCTTCCAAGG GCGTGAAGGC CGTCTGAAAA
 551  TTTTGGTGGT CGGCGGCAGT TTGGGCGCGG ACGTTTTGAA CAAAACCGTA
 601  CCGCAGGCAT TGGCTTTGCT GCCCGACAAT GCGCGTCCGC AGATGTACCA
 651  CCAATCGGGA CGGGGCAAGC TGGGCAGCTT GCAGGCGGAT TACGACGCGC
 701  TGGGCGTGCA AGCGGAATGC GTGGAATTTA TTACCGATAT GGTGTCCGCC
 751  TACCGCGATG CCGATTTGGT GATTTGCCGT GCCGGCGCGC TGACGATTGC
 801  CGAGTTGACG GCGGCGGGAT TGGGTGCGTT GTTAGTGCCG TATCCTCACG
 851  CCGTTGATGA CCATCAAACC GCCAACGCGC GTTTTATGGT GCAGGCGGAG
 901  GCGGGATTGC TGTTGCCGCA AACCCAGTTG ACGGCGGAAA AACTCGCCGA
 951  GATTCTCGGC GGCTTAAACC GCGAAAAATG CCTCAAATGG GCAGAAAACG
1001  CCCGTACGTT GGCACTGCCG CACAGTGCGG ACGACGTTGC CGAAGCCGCG
1051  ATTGCGTGTG CGGCGTAA
```

This corresponds to the amino acid sequence <SEQ ID 282; ORF 087.a>:

a087.pep

```
   1  MGGKTFMLMA GGTGGHIFPA LAVADSLRAR GHHVIWLGSK DSMEERIVPQ
  51  YDILLETLAI KGVRGNGIKR KLMLPFTLYQ TVREAQQIIR KHRVECVIGF
 101  GGFVTFPGGL AAKLLGVPIV IHEQNAVAGL SNRHLSRWAK RVLYAFPKAF
 151  SHEGGLVGNP VRADISNLPV PAERFQGREG RLKILVVGGS LGADVLNKTV
 201  PQALALLPDN ARPQMYHQSG RGKLGSLQAD YDALGVQAEC VEFITDMVSA
 251  YRDADLVICR AGALTIAELT AAGLGALLVP YPHAVDDHQT ANARFMVQAE
 301  AGLLLPQTQL TAEKLAEILG GLNREKCLKW AENARTLALP HSADDVAEAA
 351  IACAA*
```

m087/a087 85.4% identity over a 355 aa overlap

```
                   10        20        30        40        50        60
m087.pep   MGGKTFMLXXGGTGGHIFPALAVADSLRARGHHVIWLGSKDSMEERIVPQYGIRLETLAI
           ||||||||  ||||||||||||||||||||||||||||||||||||||||| | ||||||
a087       MGGKTFMLMAGGTGGHIFPALAVADSLRARGHHVIWLGSKDSMEERIVPQYDILLETLAI
                   10        20        30        40        50        60
```

```
                  70        80        90       100       110       120
m087.pep  KGVRGNGIKRKLMLPVTLYQTVREAQRIIRKHRVECVIGFGGFVTFPGGLAAKLLXVPIV
          ||||||||||||||| |||||||||:|||||||||||||||||||||||||||||| ||||
a087      KGVRGNGIKRKLMLPFTLYQTVREAQQIIRKHRVECVIGFGGFVTFPGGLAAKLLGVPIV
                  70        80        90       100       110       120

                 130       140       150       160       170       180
m087.pep  IHEQNAVAGLSNRHLSRWAKRVLYAFPKAFSHEGGLVGNPVRADISNLPVPAERFQGREG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a087      IHEQNAVAGLSNRHLSRWAKRVLYAFPKAFSHEGGLVGNPVRADISNLPVPAERFQGREG
                 130       140       150       160       170       180

                 190       200       210       220       230       240
m087.pep  RLKILVVGGSLGADVLNKTVPHALALLPDNARPHMYHQSGRGKLGILQAXXXXXXXXXXX
          |||||||||||||||||||||:||||||||||||:|||||||||||| |||
a087      RLKILVVGGSLGADVLNKTVPQALALLPDNARPQMYHQSGRGKLGSLQADYDALGVQAEC
                 190       200       210       220       230       240

                                     250       260       270       280
m087.pep  XX-------------------XXXXXXXXXAGLGALLVPYPHAVDDHQTANARFMVQAE
                                  :    : ||||||||||||||||||||||||||||||
a087      VEFITDMVSAYRDADLVICRAGALTIAELTAAGLGALLVPYPHAVDDHQTANARFMVQAE
                 250       260       270       280       290       300

                 290       300       310       320       330
m087.pep  AGLLLPQTQLTAEKLAEILGGLNREKCLKWAENARTLALPHSADDVAEAAIACAAX
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a087      AGLLLPQTQLTAEKLAEILGGLNREKCLKWAENARTLALPHSADDVAEAAIACAAX
                 310       320       330       340       350
```

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 283>:

```
g088.seq
     1  ATGTTTTTAT GGCTCGCACA TTTCAGCAAC TGGTTAACCG GTCTGAATAT
    51  TTTTCAATAC ACCACATTCC GCGCCGTTAT GGCGGCGTTG ACCGCCTTGG
   101  CGTTTTCCCT GATGTTCGGC CCGTGGACGA TACGCAGGCT GACCGCGCTC
   151  AAATGCGGGC AGGCAGTGCG TACCGACGGC CCGCAAACCC ACCTCGTCAA
   201  AAACGGCACG CCGACGATGG GCGGTTCGCT GATTCTGACC GCCATTACCG
   251  TGTCCACCCT GTTGTGGGGC AACTGGGCGA ACCCGTATAT CTGGATTCTC
   301  TTGGGCGTAC TGCTTGCCAC CGGTGCGCTC GGTTTTTACG ACGACTGGCG
   351  CAAAGTCGTT TATAAAGACC CCAACGGCGT GTCCGCCAAA TTCAAAATGG
   401  TGTGGCAGTC AAGCGTTGCC GTTatcgcCG GTttggcaTT GTTTTACctt
   451  gCcgcCAATT CCGCCAACAA TATTTTGATT GTCCCGtttT TCAAACAAAT
   501  CGCCCTGCCG CTGGGCGTGG TCGGCTTttt gGtgttgTCT TACCTGACCA
   551  TCGTCGGCAC ATCCAACGCC GTCAACCTCA CcgaCGGCTT GGACGGCCTT
   601  GCCGCcttcc cgttcgtcct cgttgccgcC GGGCTCGCCA ttttcgccTA
   651  CGTCAGCGGA CACTACCAAT TTTCCCAATA CCTCCAGCTT CCCTATGTCG
   701  CCGGCGCGAA CGAAGTCGCT ATATTCTGCA CCGCCATGTG CGGCGCGTGC
   751  CTCGGATTTT TGTGGTTCAA CGCCTATCCC GCGCAAGTCT TTATGGGCGA
   801  TGTCGGCGCG CTGGCATTGG GTGCCGCGCT CGGTaccGtt gCCGTcaTcg
   851  tCCGCCAAGA ATTTGTcctc gtcattaTGG GCGGTCTGTT cgtcgtagaa
   901  gccgtgTCCG TTATGCTTCa tgtcggCTGG TACAAGAAAA Ccaaaaaacg
   951  CATCTTcCTg acgGcaccga ttcatcacca ttaCCaactt cgatgCTGGa
  1001  aagaaacgca agtcgtcgtc CGTTtCTGGA TTAtTAccat cgtcgtggtt
  1051  tTgataggtt tGagtacccT caAAattcgc ggaaactatg ccgTCCGAAC
  1101  ACCTTTCAGA CGGCATTTGA ACGCGCAATA A
```

This corresponds to the amino acid sequence <SEQ ID 284; ORF 088.ng>:

```
g088.pep
     1  MFLWLAHFSN WLTGLNIFQY TTFRAVMAAL TALAFSLMFG PWTIRRLTAL
    51  KCGQAVRTDG PQTHLVKNGT PTMGGSLILT AITVSTLLWG NWANPYIWIL
   101  LGVLLATGAL GFYDDWRKVV YKDPNGVSAK FKMVWQSSVA VIAGLALFYL
   151  AANSANNILI VPFFKQIALP LGVVGFLVLS YLTIVGTSNA VNLTDGLDGL
```

```
201  AAFPFVLVAA GLAIFAYVSG HYQFSQYLQL PYVAGANEVA IFCTAMCGAC
251  LGFLWFNAYP AQVFMGDVGA LALGAALGTV AVIVRQEFVL VIMGGLFVVE
301  AVSVMLHVGW YKKTKKRIFL TAPIHHHYQL RCWKETQVVV RFWIITIVVV
351  LIGLSTLKIR GNYAVRTPFR RHLNAQ*
```

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 285>:

m088.seq

```
   1  ATGTTTTTAT GGCTCGCACA TTTCAGCAnC TGGTTAACCG GTCTGAATnn
  51  nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 101  nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 151  nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 201  nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 251  nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 301  nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 351  nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 401  nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 451  nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 501  nnnnnnnnnn nnnGGCGTGG TCGGCTTTTT GGTGTTGTCT TACCTGACCA
 551  TCGTCGGCAC ATCCAATGCC GTCAACCTCA CCGACGGCTT GGACGGCCTT
 601  GCGACCTTCC CCGTCGTCCT CGTTGCCGCC GGCCTCGCCA TCTTCGCCTA
 651  TGCCAGCGGC CACTCACAAT TTGCCCAATA CCTGCAATTA CCTTACGTTG
 701  CCGGCGCAAA CGAAGTGGTG ATTTTCTGTA CCGCCATGTG CGGCGCGTGC
 751  CTCGGTTTCT TGTGGTTTAA CGCCTATCCC GCGCAAGTCT TTATGGGCGA
 801  TGTCGGTGCA TTGGCATTGG GTGCCGCGCT CGGTACCGTC GCCGTTATCG
 851  TCCGCCAAGA GTTTGTCCTC GTCATTATGG GCGGATTATT TGTCGTAGAA
 901  GCCGTATCCG TTATGCTTCA GGTTGGCTGG TATAAGAAAA CCAAAAAACG
 951  CATCTTCCTG ATGGCGCCCA TCCATCACCA CTACGAACAA AAAGGCTGGA
1001  AAGAAACCCA AGTCGTCGTC CGCTTTTGGA TTATTACCAT CGTCTTGGTG
1051  TTGATCGGTT TGAGTACCCT CAAAATCCGC TGAACCTATG CCGTCTGAAC
1101  ATCTTTCAGA CGGCATTTGA ACGCGCAATA A
```

```
   1  MFLWLAHFSN WLTGLNIFQY TTFRAVMAAL TALAFSLMFG PWTIRRLTAL
  51  KCGQAVRTDG PQTHLVKNGT PTMGGSLILT AITVSTLLWG NWANPYIWIL
 101  LGVLLATGAL GFYDDWRKVV YKDPNGVSAK FKMVWQSSVA VIAGLALFYL
 151  AANSANNILI VPFFKQIALP LGVVGFLVLS YLTIVGTSNA VNLTDGLDGL
 201  AAFPFVLVAA GLAIFAYVSG HYQFSQYLQL PYVAGANEVA IFCTAMCGAC
 251  LGFLWFNAYP AQVFMGDVGA LALGAALGTV AVIVRQEFVL VIMGGLFVVE
 301  AVSVMLHVGW YKKTKKRIFL TAPIHHHYQL RCWKETQVVV RFWIITIVVV
 351  LIGLSTLKIR GNYAVRTPFR RHLNAQ*
```

This corresponds to the amino acid sequence <SEQ ID 286; ORF 088>:

m088.pep

```
   1  MFLWLAHFSX WLTGLNXXXX XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX
  51  XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX
 101  XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX
 151  XXXXXXXXXX XXXXXXXXXX XGVVGFLVLS YLTIVGTSNA VNLTDGLDGL
 201  ATFPVVLVAA GLAIFAYASG HSQFAQYLQL PYVAGANEVV IFCTAMCGAC
 251  LGFLWFNAYP AQVFMGDVGA LALGAALGTV AVIVRQEFVL VIMGGLFVVE
 301  AVSVMLQVGW YKKTKKRIFL MAPIHHHYEQ KGWKETQVVV RFWIITIVLV
 351  LIGLSTLKIR XTYAVXTSFR RHLNAQ*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 088 shows 91.7% identity over a 205 aa overlap with a predicted ORF (ORF 088.ng) from *N. gonorrhoeae:*

m088/g088

```
                                                    10        20        30
m088.pep                                  GVVGFLVLSYLTIVGTSNAVNLTDGLDGLA
                                          ||||||||||||||||||||||||||||||
g088          IAGLALFYLAANSANNILIVPFFKQIALPLGVVGFLVLSYLTIVGTSNAVNLTDGLDGLA
                    150       160       170       180       190       200
```

```
                 40        50        60        70        80        90
m088.pep    TFPVVLVAAGLAIFAYASGHSQFAQYLQLPYVAGANEVVIFCTAMCGACLGFLWFNAYPA
            :|| |||||||||||:||| ||:|||||||||||||:||||||||||||||||||||
g088        AFPFVLVAAGLAIFAYVSGHYQFSQYLQLPYVAGANEVAIFCTAMCGACLGFLWFNAYPA
                210       220       230       240       250       260

                100       110       120       130       140       150
m088.pep    QVFMGDVGALALGAALGTVAVIVRQEFVLVIMGGLFVVEAVSVMLQVGWYKKTKKRIFLM
            ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
g088        QVFMGDVGALALGAALGTVAVIVRQEFVLVIMGGLFVVEAVSVMLHVGWYKKTKKRIFLT
                270       280       290       300       310       320

                160       170       180       190       200
m088.pep    APIHHHYEQKGWKETQVVVRFWIITIVLVLIGLSTLKIRXTYAVXTSFRRHLNAQX
            ||||||: : |||||||||||||||:||||||||||| :||| | ||||||||
g088        APIHHHYQLRCWKETQVVVRFWIITIVVVLIGLSTLKIRGNYAVRTPFRRHLNAQX
                330       340       350       360       370
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 287>:

a088.seq

```
   1  ATGTTTTTAT GGCTCGCACA TTTCAGCAAC TGGTTAACCG GTCTGAATAT
  51  TTTTCAATAC ACCACATTCC GCGCCGTCAT GGCGGCGTTG ACCGCCTTGG
 101  CGTTTTCCCT GATGTTCGGC CCGTGGACGA TACGCAGGCT GACCGCGCTC
 151  AAATGCGGGC AGGCAGTGCG TACCGACGGT CCGCAAACCC ACCTCGTCAA
 201  AAACGGCACG CCGACGATGG GCGGTTCGCT GATTCTGACC GCCATTACCG
 251  TGTCCACCCT GTTGTGGGGC AACTGGGCAA ACCCGTATAT CTGGATTCTC
 301  TTGGGCGTAT TGCTCGCCAC GGGCGCACTC GGTTTTTACG ACGACTGGCG
 351  CAAAGTCGTC TATAAAGACC CCAACGGCGT GTCCGCCAAA TTCAAAATGG
 401  TGTGGCAGTC AAGCGTTGCC ATTATCGCCG GTTTGGCATT GTTTTACCTT
 451  GCCGCCAATT CCGCCAACAA TATTTTGATT GTCCCGTTCT TCAAACAAAT
 501  CGCCCTGCCG CTGGGCGTGG TCGGCTTTTT GGTGTTGTCT TACCTGACCA
 551  TCGTCGGCAC ATCCAATGCC GTCAACCTCA CCGACGGCTT GGACGGCCTT
 601  GCGACCTTCC CCGTCGTCCT CGTTGCCGCC GGCCTCGCCA TCTTCGCCTA
 651  TGCCAGCGGC CACTCACAAT TTGCCCAATA CCTGCAATTA CCTTACGTTG
 701  CCGGCGCAAA CGAAGTGGTG ATTTTCTGTA CCGCCATGTG CGGCGCGTGC
 751  CTCGGTTTCT TGTGGTTTAA CGCCTATCCC GCGCAAGTCT TTATGGGCGA
 801  TGTCGGTGCA TTGGCATTGG GTGCCGCGCT CGGTACCGTC GCCGTCATCG
 851  TCCGCCAAGA GTTTGTCCTC GTCATTATGG GCGGATTATT TGTCGTAGAA
 901  GCCGTATCCG TTATGCTTCA GGTCGGCTGG TATAAGAAAA CCAAAAAACG
 951  CATCTTCCTG ATGGCGCCCA TCCATCACCA CTACGAACAA AAAGGCTGGA
1001  AAGAAACCCA AGTCGTCGTC CGCTTTTGGA TTATTACCAT CGTCTTGGTG
1051  TTGATCGGTT TGAGTACCCT CAAAATCCGC TGAACCTATG CCGTCTGAAC
1101  ACCTTTCAGA CGGCATTTGA ACGCGCAATA A
```

This corresponds to the amino acid sequence <SEQ ID 288; ORF 088.a>:

a088.pep

```
   1  MFLWLAHFSN WLTGLNIFQY TTFRAVMAAL TALAFSLMFG PWTIRRLTAL
  51  KCGQAVRTDG PQTHLVKNGT PTMGGSLILT AITVSTLLWG NWANPYIWIL
 101  LGVLLATGAL GFYDDWRKVV YKDPNGVSAK FKMVWQSSVA IIAGLALFYL
 151  AANSANNILI VPFFKQIALP LGVVGFLVLS YLTIVGTSNA VNLTDGLDGL
 201  ATFPVVLVAA GLAIFAYASG HSQFAQYLQL PYVAGANEVV IFCTAMCGAC
 251  LGFLWFNAYP AQVFMGDVGA LALGAALGTV AVIVRQEFVL VIMGGLFVVE
 301  AVSVMLQVGW YKKTKKRIFL MAPIHHHYEQ KGWKETQVVV RFWIITIVLV
 351  LIGLSTLKIR *TYAV*TPFR RHLNAQ*
```

m088/a088 99.5% identity over a 205 aa overlap

```
                 150       160       170       180       190       200
m088.pep    XXXXXXXXXXXXXXXXXXXXXXXXXXXXXGVVGFLVLSYLTIVGTSNAVNLTDGLDGLA
                                         |||||||||||||||||||||||||||||
a088        IAGLALFYLAANSANNILIVPFFKQIALPLGVVGFLVLSYLTIVGTSNAVNLTDGLDGLA
```

```
                    150         160         170         180         190         200

                    210         220         230         240         250         260
m088.pep    TFPVVLVAAGLAIFAYASGHSQFAQYLQLPYVAGANEVVIFCTAMCGACLGFLWFNAYPA
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a088        TFPVVLVAAGLAIFAYASGHSQFAQYLQLPYVAGANEVVIFCTAMCGACLGFLWFNAYPA
                    210         220         230         240         250         260

                    270         280         290         300         310         320
m088.pep    QVFMGDVGALALGAALGTVAVIVRQEFVLVIMGGLFVVEAVSVMLQVGWYKKTKKRIFLM
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a088        QVFMGDVGALALGAALGTVAVIVRQEFVLVIMGGLFVVEAVSVMLQVGWYKKTKKRIFLM
                    270         280         290         300         310         320

                    330         340         350         360         370
m088.pep    APIHHHYEQKGWKETQVVVRFWIITIVLVLIGLSTLKIRXTYAVXTSFRRHLNAQX
            |||||||||||||||||||||||||||||||||||||||||||||| |||||||||
a088        APIHHHYEQKGWKETQVVVRFWIITIVLVLIGLSTLKIRXTYAVXTPFRRHLNAQX
                    330         340         350         360         370
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 289>:

```
g089.seq
      1  ATGCCGCCCA AAATCACGAA GAGCGGGTTT TGCAAACCGG CAATCGCGGC
     51  GGCGGTCGCG CCGACATTCG TGCCTTTGCT GTCGTCGATG AATACCACGC
    101  CGTTTTTCTC GCCGATTTTT TCCACACGGT GCGGCAAGCC TTGGAAGGTT
    151  TTGACGTGTT CCAGCAATGC TTCGCGCGGC AAACCGACGG CCTCGCACAA
    201  AGCCACGGCA GCCATAACGT TGGCGGCGTT GTGCAAACCT TGCAGCGGGA
    251  TGTCTTGCGT AGAAATCAAA TCTTCATTGC CTTGTTTTAA ACAGCCCGTC
    301  CCGCGTTCCA ACCAAAAATC GGCTTCGTGT TCCAAGGAAA ACCGTTTCAC
    351  TTCACGCCCT GCCCGTTTCA TGGCGCGGCA GAACACGTCG TCCGCATTCA
    401  AAACCTGCAC TCCATCGCCA CGGAAAATCT CGGCTTTGGT ATGCGCGTAG
```

This corresponds to the amino acid sequence <SEQ ID 290; ORF 089.ng>:

```
g089.pep
      1  MPPKITKSGF CKPAIAAAVA PTFVPLLSSM NTTPFFSPIF STRCGKPWKV
     51  LTCSSNASRG KPTASHKATA AITLAALCKP CSGMSCVEIK SSLPCFKQPV
    101  PRSNQKSASC SKENRFTSRP ARFMARQNTS SAFKTCTPSP RKISALVCA*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 291>:

```
m089.seq
      1  ATGCCGCCCA AAATCACkAw GAGCGGATTT TGCAAACCGG CAATCGCGGC
     51  GGCAGTCGCG CCGACATTCG TGCCTTTGCT GTCGTCGATA AACACCACGC
    101  CGTTTTTCTC GCCGATTTTT TCCACGCGGT GCGGCAGGCC TTGGAAGGTT
    151  TTGACGTGTT CGAGCAATGC TTCGCGCGAC AAACCGATGG CCTCACACAA
    201  AGCCACkGCA GCCATGACGT TAGCGGCGTT GTGCAkACCT TGCAACGGwA
    251  TGTCTTGCGT GACAATCAAA TCTTCATTGC CTTGTTTCAG GCGGCCTGTC
    301  TCGCGTTCCA ACCAGAAATC AGCTTCGTGT TCCAACGAAA ACCATTTTAC
    351  CTCGCGCCCG GCACGCTTCA TCGCGCGGCA GAACGCATCG TCCGCATTCA
    401  AAACCTGCAC GCCGTCGCCA CGGAAAATCT TGGCTTTGGT ATGCGCATAG
```

This corresponds to the amino acid sequence <SEQ ID 292; ORF 089>:

```
m089.pep
      1  MPPKITXSGF CKPAIAAAVA PTFVPLLSSI NTTPFFSPIF STRCGRPWKV
     51  LTCSSNASRD KPMASHKATA AMTLAALCXP CNGMSCVTIK SSLPCFRRPV
    101  SRSNQKSASC SNENHFTSRP ARFIARQNAS SAFKTCTPSP RKILALVCA*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 089 shows 88.6% identity over a 149 aa overlap with a predicted ORF (ORF 089.ng) from *N. gonorrhoeae:*

m089/g089

```
                10        20        30        40        50        60
m089.pep   MPPKITXSGFCKPAIAAAVAPTFVPLLSSINTTPFFSPIFSTRCGRPWKVLTCSSNASRD
           |||||| |||||||||||||||||||||||:||||||||||||||:||||||||||||
g089       MPPKITKSGFCKPAIAAAVAPTFVPLLSSMNTTPFFSPIFSTRCGKPWKVLTCSSNASRG
                10        20        30        40        50        60

                70        80        90       100       110       120
m089.pep   KPMASHKATAAMTLAALCXPCNGMSCVTIKSSLPCFRRPVSRSNQKSASCSNENHFTSRP
           || ||||||||:|||||| ||:||||| ||||||||::|| |||||||||:||:|||||
g089       KPTASHKATAAITLAALCKPCSGMSCVEIKSSLPCFKQPVPRSNQKSASCSKENRFTSRP
                70        80        90       100       110       120

               130       140       150
m089.pep   ARFIARQNASSAFKTCTPSPRKILALVCAX
           |||:||||:|||||||||||||| ||||||
g089       ARFMARQNTSSAFKTCTPSPRKISALVCAX
               130       140       150
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 293>:

a089.seq

```
  1  ATGCCGCCTA AAATCACGAA GAGCGGATTT TGCAAACCGG CAATCGCGGC
 51  GGCGGTCGCA CCGACGTTCG TGCCTTTGCT GTCGTCGATG AACACCACGC
101  CATTTTTCTC GCCGATTTTT TCCACGCGGT GCGGCAGGCC TTGAAAGGTT
151  TTGACGTGTT CGAGCAATGC TTCGCGCGGC AAACCGACGG CTTCGCACAA
201  GGCAACGGCA GCCATCACGT TAGTGGCGTT GTGCAAGCCT TGCAGCGGAA
251  TATCTTGCGT GGCAATCAAA TCTTCATTGC CTTGTTTCAG GCGACCTGTC
301  TCACGTTCCA ACCAAAAATC GGCTTCGTAT TCCAACGAAA ACCATTTCAC
351  CTCGCGCCCG GCGCGCTTCA TCGCACGACA GAACGCATCG TCCGCATTCA
401  AAACCTGCAC ACCGTCGCCA CGGAAAATCT TGGCTTTGGT ATGCGCGTAG
```

This corresponds to the amino acid sequence <SEQ ID 294; ORF 089.a>:

a089.pep

```
  1  MPPKITKSGF CKPAIAAAVA PTFVPLLSSM NTTPFFSPIF STRCGRP*KV
 51  LTCSSNASRG KPTASHKATA AITLVALCKP CSGISCVAIK SSLPCFRRPV
101  SRSNQKSASY SNENHFTSRP ARFIARQNAS SAFKTCTPSP RKILALVCA*
```

m089/a089 91.9% identity over a 149 aa overlap

```
                10        20        30        40        50        60
m089.pep   MPPKITXSGFCKPAIAAAVAPTFVPLLSSINTTPFFSPIFSTRCGRPWKVLTCSSNASRD
           |||||| |||||||||||||||||||||||:||||||||||||||||| ||||||||||
a089       MPPKITKSGFCKPAIAAAVAPTFVPLLSSMNTTPFFSPIFSTRCGRPXKVLTCSSNASRG
                10        20        30        40        50        60

                70        80        90       100       110       120
m089.pep   KPMASHKATAAMTLAALCXPCNGMSCVTIKSSLPCFRRPVSRSNQKSASCSNENHFTSRP
           || ||||||||:||:||| ||:|:|||:||||||||||||||||||||| ||||||||||
a089       KPTASHKATAAITLVALCKPCSGISCVAIKSSLPCFRRPVSRSNQKSASYSNENHFTSRP
                70        80        90       100       110       120

               130       140       150
m089.pep   ARFIARQNASSAFKTCTPSPRKILALVCAX
           ||||||||||||||||||||||||||||||
a089       ARFIARQNASSAFKTCTPSPRKILALVCAX
               130       140       150
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 295>:

g090.seq

```
  1  ATGCGCGTAG TCGAGCAAAT CGTCGTAGCG GTCGAGATGG TCTTCGGAAA
```

```
 51  TGTTCATCAC CGTCGCCGCA GTCGGGCGCA GGCTTTCGGT GTTTTCCAGT
101  TGGAAGCTGG AAAGCTCcca CACCCACACG TCCGCCTTTT TGCCTTCgcg
151  ctgCAATtct gcctccaaga cgggcgtacc gatATTGCCC GCAATGAcgg
201  tatccagccc gcacttgatg CAGAGatagc ggaccaggct ggttaccgTG
251  GTTttgccgt tgctgCcggt aatcgCaatc accttgtcgC CGCGGCGGtt
301  cAcaaTGTCc gccaGCAATt ggATGTCGCC TAgCACGCGC .ccgccgTTT
351  TGCttga
```

This corresponds to the amino acid sequence <SEQ ID 296; ORF 090.ng>:

```
g090.pep
  1  MRVVEQIVVA VEMVFGNVHH RRRSRAQAFG VFQLEAGKLP HPHVRLFAFA
 51  LQFCLQDGRT DIARNDGIQP ALDAEIADQA GYRGFAVAAG NRNHLVAAAV
101  HNVRQQLDVA XHAXRRFA*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 297>:

```
m090.seq
  1  ATGCGCATAG TCGAGCAAGT CGTCGTAGCG GTCGAGATGG TCTTCGGAAA
 51  TGTTCAGCAC CGTCGCCGCA GTCGGACGCA GGCTTTCGGT GTTTTCCAGT
101  TGGAAGCTGG AAAGCTCCAA CACCCACACG TCCGCCTTTT TGCCTTCGCG
151  CTGCCATTCC GCCTCCAAAA CCGGCGTGCC GATATTGCCC GCGATAACGG
201  TATCCAGCCC GCACTTGATA CAGAGATAGC CGACCAGGCT CGTTACCGTG
251  GTTTTGCCGT TGCTGCCGGT AATCGCAATT ACCTTGTCGT CCCGGCGGTT
301  CACAATGTCC GCCAGCAATT CGATGTCGCC CAACACGCGT .CCGCCGTTT
351  TGCTTGA
```

This corresponds to the amino acid sequence <SEQ ID 298; ORF 090>:

```
m090.pep
  1  MRIVEQVVVA VEMVFGNVQH RRRSRTQAFG VFQLEAGKLQ HPHVRLFAFA
 51  LPFRLQNRRA DIARDNGIQP ALDTEIADQA RYRGFAVAAG NRNYLVVPAV
101  HNVRQQFDVA QHAXRRFA*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 090 shows 83.9% identity over a 118 aa overlap with a predicted ORF (ORF 090.ng) from *N. gonorrhoeae:*

```
m090/g090

                        10        20        30        40        50        60
m090.pep       MRIVEQVVVAVEMVFGNVQHRRRSRTQAFGVFQLEAGKLQHPHVRLFAFALPFRLQNRRA
               ||:|||:|||||||||||:||||||:||||||||||||| ||||||||||| | ||: |:
g090           MRVVEQIVVAVEMVFGNVHHRRRSRAQAFGVFQLEAGKLPHPHVRLFAFALQFCLQDGRT
                        10        20        30        40        50        60

                        70        80        90       100       110       119
m090.pep       DIARDNGIQPALDTEIADQARYRGFAVAAGNRNYLVVPAVHNVRQQFDVAQHAXRRFAX
               ||||::|||||||:|||||| ||||||||||||:||: ||||||||:||| ||||||||
g090           DIARNDGIQPALDAEIADQAGYRGFAVAAGNRNHLVAAAVHNVRQQLDVAXHAXRRFAX
                        70        80        90       100       110
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 299>:

```
a090.seq
  1  ATGCGCGTAG TCGAGCAAGT CGTCGTAGCG GTCGAGATGG TCTTCGGAAA
 51  TGTTCAGCAC TGTCGCCGCA GTCGGGCGCA GGCTTTCGGT GTTTTCCAGT
101  TGGAAACTGG AAAGCTCCAA CACCCACACG TCCGCCTTTT TGCCTTCGCG
151  CTGCAATTCC GCCTCCAAAA CCGGCGCGCC GATATTGCCC GCGATAACGG
201  TATCCAGCCC ACACTTGATG CAGAGATAGC CGACCAGGCT CGTTACCGTG
251  GTTTTGCCGT TGCTGCCGGT AATCGCAATC ACCTTGTCGC CGCGGCGGTT
301  CACAATGTCC GCCAGCAATT CGATGTCGCC CAACACGCGT C.CGCCGTTT
351  CGCTTAA
```

This corresponds to the amino acid sequence <SEQ ID 300; ORF 090.a>:

a090.pep

```
  1  MRVVEQVVVA VEMVFGNVQH CRRSRAQAFG VFQLETGKLQ HPHVRLFAFA
 51  LQFRLQNRRA DIARDNGIQP TLDAEIADQA RYRGFAVAAG NRNHLVAAAV
101  HNVRQQFDVA QHAXRRFA*
```

m09/a090 91.5% identity over a 117 aa overlap

```
                    10        20        30        40        50        60
m090.pep    MRIVEQVVVAVEMVFGNVQHRRRSRTQAFGVFQLEAGKLQHPHVRLFAFALPFRLQNRRA
            ||:||||||||||||||||| ||||:|||||||||:||||||||||||||| ||||||||
a090        MRVVEQVVVAVEMVFGNVQHCRRSRAQAFGVFQLETGKLQHPHVRLFAFALQFRLQNRRA
                    10        20        30        40        50        60

                    70        80        90       100       110       119
m090.pep    DIARDNGIQPALDTEIADQARYRGFAVAAGNRNYLVVPAVHNVRQQFDVAQHAXRRFAX
            ||||||||||:||:||||||||||||||||||||:||: ||||||||||||||||||||
a090        DIARDNGIQPTLDAEIADQARYRGFAVAAGNRNHLVAAAVHNVRQQFDVAQHAXRRFAX
                    70        80        90       100       110
```

The following partial DNA sequence was identified in *N. gonorrhoeae*

g090-1.seq This sequence contains multiple stop codons (not shown)

This corresponds to the amino acid sequence < ORF 090-1.ng>:

g090-1.pep (not shown)

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 2>:

m090-1.seq

```
   1  ATGACGGCGT TTGCATTTCA GACGGCATCA CAAAGCCTTA AACGCTTCGA
  51  TAAACACTTC CGAACGGTGC GCGTAGCCTT TGAACATATC AAAGCTCGCG
 101  CAGGCGGGGC TGAGCAACAC AATATCGCCT GCTTCGGCTT GGGCATATGC
 151  CGTCTGAACG GCTTCTCCCA AAGTGGCGCA GTCGGTCATA TTCAAGCCGC
 201  AGCCGTCCAA ATCGCGGCGG ATTTGCGGCG CATCGACACC AATCAAGAAC
 251  ACGCCTTTTG CCTTGCCTAC CAGTGCATCG CGCAGGGGCG TGAAGTCCTG
 301  CCCTTTACCC ATGCCGCCCA AAATCACGAA GAGCGGATTT TGCAAACCGG
 351  CAATCGCGGC GGCAGTCGCG CCGACATTCG TGCCTTTGCT GTCGTCGATA
 401  AACACCACGC CGTTTTTCTC GCCGATTTTT TCCACGCGGT GCGGCAGGCC
 451  TTGGAAGGTT TTGACGTGTT CGAGCAATGC TTCGCGCGAC AAACCGATGG
 501  CCTCACACAA AGCCACGGCA GCCATGACGT TAGCGGCGTT GTGCAGACCT
 551  TGCAACGGAA TGTCTTGCGT GACAATCAAA TCTTCATTGC CTTGTTTCAG
 601  GCGGCCTGTC TCGCGTTCCA ACCAGAAATC AGCTTCGTGT TCCAACGAAA
 651  ACCATTTTAC CTCGCGCCCG GCACGCTTCA TCGCGCGGCA GAACGCATCG
 701  TCCGCATTCA AAACCTGCAC GCCGTCGCCA CGGAAAATCT TGGCTTTGGT
 751  ATGCGCATAG TCGAGCAAGT CGTCGTAGCG GTCGAGATGG TCTTCGGAAA
 801  TGTTCAGCAC CGTCGCCGCA GTCGGACGCA GGCTTTCGGT GTTTTCCAGT
 851  TGGAAGCTGG AAAGCTCCAA CACCCACACG TCCGCCTTTT TGCCTTCGCG
 901  CTGCCATTCC GCCTCCAAAA CCGGCGTGCC GATATTGCCC GCGATAACGG
 951  TATCCAGCCC GCACTTGATA CAGAGATAGC CGACCAGGCT CGTTACCGTG
1001  GTTTTGCCGT TGCTGCCGGT AATCGCAATT ACCTTGTCGT CCCGGCGGTT
1051  CACAATGTCC GCCAGCAATT CGATGTCGCC CAACACGCGT CCGCCGTTTT
1101  GCTTGAACGC CTCAATATCC GGCTGCCGCT CGCTGATGCC GGGACTGAGA
1151  GCCAGAATAT CGAAACCGTT GTCCAGCGCA TCTTTCAGAC GGCCCGTGTA
1201  AAACACCAAC CCGTCAAACA TCTTACCGAT TTGCGACACG CGTTCCGGCT
1251  TCAGCTCCGC ATCATACGCA GCAACCTCCG CGCCGTTTTT GCGCAGGTAG
1301  GCAATCATGG AAATACCCGT ACCGCCGAGT CCGGCGACGA GGATTTTTTT
1351  GTTTTGAAAA GTCATTTTGG TTTGTCCTAA
```

This corresponds to the amino acid sequence <SEQ ID 3; ORF 090-1>:

m090-1.pep

```
  1  MTAFAFQTAS QSLKRFDKHF RTVRVAFEHI KARAGGAEQH NIACFGLGIC
 51  RLNGFSQSGA VGHIQAAAVQ IAADLRRIDT NQEHAFCLAY QCIAQGREVL
101  PFTHAAQNHE ERILQTGNRG GSRADIRAFA VVDKHHAVFL ADFFHAVRQA
151  LEGFDVFEQC FARQTDGLTQ SHGSHDVSGV VQTLQRNVLR DNQIFIALFQ
201  AACLAFQPEI SFVFQRKPFY LAPGTLHRAA ERIVRIQNLH AVATENLGFG
251  MRIVEQVVVA VEMVFGNVQH RRRSRTQAFG VFQLEAGKLQ HPHVRLFAFA
301  LPFRLQNRRA DIARDNGIQP ALDTEIADQA RYRGFAVAAG NRNYLVVPAV
351  HNVRQQFDVA QHASAVLLER LNIRLPLADA GTESQNIETV VQRIFQTARV
401  KHQPVKHLTD LRHAFRLQLR IIRSNLRAVF AQVGNHGNTR TAESGDEDFF
451  VLKSHFGLS*
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 303>:

```
g091.seq
      1  ATGGAAATAC CCGTGCCGCC AAGTCCGGCG ACGAGGATTT TTTTGTTTGA
     51  AAGTCATTTT GGTTTTGTCC TAAAACAAAT CATATTGGGC AGGAGACGTC
    101  CGCCCTTGCC CAAGCCGCTT TCAGACGGCA TCGCGAGCCG ATTAATAACC
    151  CGCCTTCAGG CGTTGGTCAT TGTCGCAGCT GTTTTGGTCT CCGTTTTGAC
    201  AAGCCTTGCC AAGCCATTGT TGAGCGAGCG CAAGGTCTTG GCGCACGCCG
    251  CGTCCATCGT AATACATCAA GCCCAAATTG TATTGGGCTT GGGCATCCCC
    301  TTGTTCTGA
```

This corresponds to the amino acid sequence <SEQ ID 304; ORF 091.ng>:

```
g091.pep
      1  MEIPVPPSPA TRIFLFESHF GFVLKQIILG RRRPPLPKPL SDGIASRLIT
     51  RLQALVIVAA VLVSVLTSLA KPLLSERKVL AHAASIVIHQ AQIVLGLGIP
    101  LF*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 305>:

```
m091.seq
      1  ATGGAAATAC CCGTACCGCC GAGTCCGGCG ACGAGGATTT TTTTGTTTGA
     51  AAAGTCATTT TGGTTTGTCC TAAAACAAAT CATATTGAGC AGGAGATGTC
    101  CGCCCCTGCC CAAGCCGCTT TCAGACGGCA TCGCGAGCTG TTCAATAACC
    151  CGCCTTCAGG CGTTGGTCAT TGTCGCAGCC GTCTTGGTCT CCGTTTTGAC
    201  AAGCCTTGCC AAACCATTCT TGTGCAAGGG CGCGGTCTTG GCGCACGCCG
    251  CGTCTTTCGG CATACATCAC GCCCAAATTG TTTTGGGCTT GGGCTACCCC
    301  CTGCGC...
```

This corresponds to the amino acid sequence <SEQ ID 306; ORF 091>:

```
m091.pep
      1  MEIPVPPSPA TRIFLFEKSF WFVLKQIILS RRCPPLPKPL SDGIASCSIT
     51  RLQALVIVAA VLVSVLTSLA KPFLCKGAVL AHAASFGIHH AQIVLGLGYP
    101  LR.
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 091 shows 84.2% identity over a 101 aa overlap with a predicted ORF (ORF 091.ng) from *N. gonorrhoeae:*

```
m091/g091

                         10        20        30        40        50        60
m091.pep         MEIPVPPSPATRIFLFEKSFWFVLKQIILSRRCPPLPKPLSDGIASCSITRLQALVIVAA
                 |||||||||||||||||: | ||||||||:|| |||||||||||||  ||||||||||||
g091             MEIPVPPSPATRIFLFESHFGFVLKQIILGRRRPPLPKPLSDGIASRLITRLQALVIVAA
                         10        20        30        40        50        60

                         70        80        90       100
m091.pep         VLVSVLTSLAKPFLCKGAVLAHAASFGIHHAQIVLGLGYPLR
                 ||||||||||||:| :  |||||||: ||:|||||||| ||
g091             VLVSVLTSLAKPLLSERKVLAHAASIVIHQAQIVLGLGIPLFX
                         70        80        90       100
```

The following partial DNA sequence was identified in *N. meningiditis* <SEQ ID 307>:

```
a091.seq
      1  ATGGAAATAC CCGTGCCGCC AAGTCCGGCG ACGAGGATTT TTTTGTTTTG
     51  GAAATCATTT TGGTTTGTCC TAAAACAAAT CATATTGAGC AGGGGATGTC
    101  TGATCCTGCT CAAGCCGCTT TCAGACGGCA TCGCGAGCTG TTCAATAACC
    151  CGCTTTCAGG CGTTGGTCAT TGTCGCAGCT GTCTTGGTAT CCGTTTTGAC
    201  AAGCCTTGCC AAGCCATTCT TGTGCAAGGG CGCGGTCTTG GCGCACGCCG
    251  CGTCTTTCGG CATACATCAC GCCCAAATTG TTTTGGGC
```

This corresponds to the amino acid sequence <SEQ ID 308; ORF 091.a>:

```
a091.pep
      1  MEIPVPPSPA TRIFLFWKSF WFVLKQIILS RGCLILLKPL SDGIASCSIT
     51  RFQALVIVAA VLVSVLTSLA KPFLCKGAVL AHAASFGIHH AQIVLG
```

m091/a091 93.8% identity over a 96 aa overlap

```
                   10        20        30        40        50        60
m091.pep    MEIPVPPSPATRIFLFEKSFWFVLKQIILSRRCPPLPKPLSDGIASCSITRLQALVIVAA
            |||||||||||||||| ||||||||||||| |   | ||||||||||||||:||||||||
a091        MEIPVPPSPATRIFLFWKSFWFVLKQIILSRGCLILLKPLSDGIASCSITRFQALVIVAA
                   10        20        30        40        50        60

                   70        80        90       100
m091.pep    VLVSVLTSLAKPFLCKGAVLAHAASFGIHHAQIVLGLGYPLR
            ||||||||||||||||||||||||||||||||||||
a091        VLVSVLTSLAKPFLCKGAVLAHAASFGIHHAQIVLG
                   70        80        90
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 309>:

```
g092.seq
      1  ATGTTTTTTA TTTCAATCCG CTATATATTT GTCAGAAAAC TATGGTGCGC
     51  AAACGGTCAG ACCTTTAAAA TAACGCCTTT ACGCACTAAA AACCAACCGG
    101  AACGCAACAT TATGATGAAA AATCGAGTAA GCAACATCCA TTTTGTCGGT
    151  ATCGGCGGCG TCGGCATGAG CGGTATCGCC GAAGTCTTGC ACAATTTGGG
    201  CTTTAAAGTT TCCGGTTCGG ATCAGGCGCG AAATGCCGCT ACCGAGCATT
    251  TGAGCAGCCT GGGCATTCAA GTTTATCCCG GCCATACCGC AGAACACGTT
    301  AACGGTgcgg ATGTCGTCGT TGCCTCTACC GCCGTCAAGA AAGAAaatcC
    351  CGAAGTtgtc gcTGCGTTGG AGCGGCAAAT TCCCGTTATT CCGCGCGCCT
    401  TGATGCTGGC AGAGCTGATG CGCTTCCGTG ACGgcatcgc cattgccggT
    451  ACGCACGGCA AAACCACGAC CACCAGCCTG ACCGCCTCCA TCCTCGGCGC
    501  GGCAGGACTC GACCCCACTT TCGTTATCGG CGGCAAACTC AACGCCGCAG
    551  GCACCAACGC CCGCTTGGGC AAAGGCGAAT ACATCGTTGC CGAAGCCGAC
    601  GAATCCGATG CCTCTTTCCT ACATCTGACC CCGATTATGT CCGTCGTTAC
    651  CAATATCGAC GAAGACCATA TGGATACCTA CGGGCACAGC GTCGAAAAAC
    701  TGCATCAGGC GTTTATCGAT TTCATCCACC GTATGCCCTT CTACGGCAAA
    751  GCCTTTTTGT GTGTTGACAG CGAACACGTC CGCGCGATTT TGCCCAAAGT
    801  GAGCAAACCT TATGCTACTT ACGGTTTGGA CGATACCGCC GACATCTACG
    851  CCACCGACAT CGAAAACGTC GGCGCGCAAA TGAAATTCAC CGTCCATGTT
    901  CAAATGAAAG GACATGAGCA GGGGTCGTTT GAAGTCGTGC TGAATATGCC
    951  CGGCAGACAC AACGTGCTGA ACGCATTGGC AGCCATCGGc gtggcGCTgg
   1001  aagtcGgCGC ATcggttgAA GCGAtcCAAA AaggCTTGCT CGGCTTTGAA
   1051  GGCGTCGGCC GCCGCTTCCA AAAATAcggc gacatCAagt tgccaaacgg
   1101  cggGaccgCT TTgctGGTGG ACGATTAcgg ACACCACCCC GTCGAAATGG
   1151  CGGcaaccct tgccgcTGCA CGCGGCGCGT ATCCGGAAAA acgtTTGGTG
   1201  CtcgCCTTCC AGCCGCACCG CTATACCCGC ACGCGCGATT TGTTTGAAGA
   1251  CTTTACCAAA GTACTCAATA CCGTTGatgC GCTGGTACTG ACCGAAGTTT
   1301  AtgccgccgG CGAAGAGCCG GTTGCCGCCG CCGactcCCG CGCCTTGGCG
   1351  CGTGCTATCC GCGTATTGGG CAAACTTGAG CCGATTTACT GCGAAAatgt
   1401  cgccgACCTG CCGCAAATGC TGATGAATGT TTTACAGGAT Ggcgatgttg
   1451  tgttgAATAT GggTgcggga agcatcaacc gcgttccttc cgcgctgttg
   1501  gaattgtcga AACAGAtttg A
```

This corresponds to the amino acid sequence <SEQ ID 310; ORF 092.ng>:

```
g092.pep
      1  MFFISIRYIF VRKLWCANGQ TFKITPLRTK NQPERNIMMK NRVSNIHFVG
     51  IGGVGMSGIA EVLHNLGFKV SGSDQARNAA TEHLSSLGIQ VYPGHTAEHV
    101  NGADVVVAST AVKKENPEVV AALERQIPVI PRALMLAELM RFRDGIAIAG
    151  THGKTTTTSL TASILGAAGL DPTFVIGGKL NAAGTNARLG KGEYIVAEAD
    201  ESDASFLHLT PIMSVVTNID EDHMDTYGHS VEKLHQAFID FIHRMPFYGK
```

```
251  AFLCVDSEHV RAILPKVSKP YATYGLDDTA DIYATDIENV GAQMKFTVHV
301  QMKGHEQGSF EVVLNMPGRH NVLNALAAIG VALEVGASVE AIQKGLLGFE
351  GVGRRFQKYG DIKLPNGGTA LLVDDYGHHP VEMAATLAAA RGAYPEKRLV
401  LAFQPHRYTR TRDLFEDFTK VLNTVDALVL TEVYAAGEEP VAAADSRALA
451  RAIRVLGKLE PIYCENVADL PQMLMNVLQD GDVVLNMGAG SINRVPSALL
501  ELSKQI*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 311>:

```
m092.seq
   1  ATGTTTTTTA TTTCAATCCG CTATATATTT GTCAGAAAAC TATGGCGCGC
  51  AAACGGTCAG CCCTTTAAAA TAACGCCTTT ACGCATCGAA AATCCACCGG
 101  AACGCAACAT TATGATGAAA AATCGAGTTA CGAACATCCA TTTTGTCGGT
 151  ATCGGCGGCG TCGGCATGAG CGGCATCGCC GAAGTCTTGC ACAATTTGGG
 201  CTTTAAAGTT TCCGGTTCGG ATCAgGCGCG AAATGCCGCT ACCGAGCATT
 251  TGGGCAGCCT GGGCATTCAA GTTTATCCCG GCCATACCGC CGAACACGTT
 301  AACGGTGCGG ATGTCGTCGT TACCTCTACC GCCGTCAAAA AAGAAAATCC
 351  CGAAGTTGTC GCTGCGTTGG AGCAGCAAAT TCCCGTTATT CCGCGCGCCC
 401  TGATGTTGGC GGAGTTGATG CGCTTCCGTG ACGGCATCGC CATTGCCGGC
 451  ACGCACGGCA AAACCACGAC CACCAGCCTG ACCGCCTCCA TCCTCGGCGC
 501  GGCAGGACTT GACCCGACTT TCGTTATCGG CGGCAAACTC AACGCCGCAG
 551  GCACTAACGC CCGCTTGGGC AAAGGCGAAT ACATCGTTGC CGAAGCCGAC
 601  GAGTCGGATG CATCCTTTCT GCACCTGACA CCGATTATGT CCGTCGTTAC
 651  CAATATCGAC GAAGACCATA TGGATACCTA CGGGCACAGC GTCGAAAAAC
 701  TGCATCAGGC GTTTATCGAT TTCATCCACC GTATGCCCTT CTACGGCAAA
 751  GCCTTTTTGT GTATTGACAG CGAACACGTC CGCGCGATTT TGCCCAAAGT
 801  GAGCAAACCT TATGCTACTT ACGGTTTGGA CGATACCGCC GACATCTACG
 851  CCACCGACAT CGAAAACGTC GGCGCGCAAA TGAAATTCAC CGTCCATGTT
 901  CAAATGAAAG GACATGAGCA GGGGTCGTTT GAAGTCGTGC TGAATATGCC
 951  CGGCAGACAC AACGTGCTGA ACGCATTGGC AGCCATCGGC GTGGCGCTGG
1001  AAGTCGGCGC ATCGGTTGAA GCGATCCAAA AAGGCTTGCT CGGCTTTGAA
1051  GGCGTCGGCC GCCGCTTCCA AAAATACGGC GACATCAAGT TGCCAAACGG
1101  CGGGACCGCG CTCTTGGTGG ACGACTACGG ACACCACCCC GTCGAAATGG
1151  CGGCGACCCT TGCCGCCGCA CGCGGCGCGT ATCTGGAAAA ACGTTTGGTA
1201  CTCGCCTTCC AGCCGCACCG CTATACCCGC ACGCGCGATT TGTTTGAAGA
1251  CTTTACCAAA GTCCTCAATA CCGTTGACGC GCTGGTGCTG ACCGAAGTTT
1301  ATGCCGCCGG TGAAGAGCCG ATTGCCGCCG CCGATTCCCG CGCTCTTGCC
1351  CGCGCCATCC GCGTGTTGGG CAAACTCGAG CCGATTTACT GCGAAAACGT
1401  TGCCGATCTG CCCGAAATGC TGTTGAACGT TTTGCAGGAC GGCGACATCG
1451  TGTTGAATAT GGGCGCGGGA AGCATCAACC GCGTCCCCGC CGCGCTGCTG
1501  GCATTGTCGA AACAGATTTG A
```

This corresponds to the amino acid sequence <SEQ ID 312; ORF 092>:

```
m092.pep
   1  MFFISIRYIF VRKLWRANGQ PFKITPLRIE NPPERNIMMK NRVTNIHFVG
  51  IGGVGMSGIA EVLHNLGFKV SGSDQARNAA TEHLGSLGIQ VYPGHTAEHV
 101  NGADVVVTST AVKKENPEVV AALEQQIPVI PRALMLAELM RFRDGIAIAG
 151  THGKTTTTSL TASILGAAGL DPTFVIGGKL NAAGTNARLG KGEYIVAEAD
 201  ESDASFLHLT PIMSVVTNID EDHMDTYGHS VEKLHQAFID FIHRMPFYGK
 251  AFLCIDSEHV RAILPKVSKP YATYGLDDTA DIYATDIENV GAQMKFTVHV
 301  QMKGHEQGSF EVVLNMPGRH NVLNALAAIG VALEVGASVE AIQKGLLGFE
 351  GVGRRFQKYG DIKLPNGGTA LLVDDYGHHP VEMAATLAAA RGAYLEKRLV
 401  LAFQPHRYTR TRDLFEDFTK VLNTVDALVL TEVYAAGEEP IAAADSRALA
 451  RAIRVLGKLE PIYCENVADL PEMLLNVLQD GDIVLNMGAG SINRVPAALL
 501  ALSKQI*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 092 shows 96.6% identity over a 506 aa overlap with a predicted ORF (ORF 092.ng) from *N. gonorrhoeae:*

```
m092/g092

                    10        20        30        40        50        60
m092.pep    MFFISIRYIFVRKLWRANGQPFKITPLRIENPPERNIMMKNRVTNIHFVGIGGVGMSGIA
```

```
                 ||||||||||||'''  |||| ||||||| :| ||||||||||:|||||||||||||||
g092             MFFISIRYIFVRKLWCANGQTFKITPLRTKNQPERNIMMKNRVSNIHFVGIGGVGMSGIA
                         10        20        30        40        50        60

                         70        80        90       100       110       120
m092.pep         EVLHNLGFKVSGSDQARNAATEHLGSLGIQVYPGHTAEHVNGADVVVTSTAVKKENPEVV
                 ||||||||||||||||||||||||:||||||||||||||||||||||:||||||||||||
g092             EVLHNLGFKVSGSDQARNAATEHLSSLGIQVYPGHTAEHVNGADVVVASTAVKKENPEVV
                         70        80        90       100       110       120

                        130       140       150       160       170       180
m092.pep         AALEQQIPVIPRALMLAELMRFRDGIAIAGTHGKTTTTSLTASILGAAGLDPTFVIGGKL
                 ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
g092             AALERQIPVIPRALMLAELMRFRDGIAIAGTHGKTTTTSLTASILGAAGLDPTFVIGGKL
                        130       140       150       160       170       180

                        190       200       210       220       230       240
m092.pep         NAAGTNARLGKGEYIVAEADESDASFLHLTPIMSVVTNIDEDHMDTYGHSVEKLHQAFID
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g092             NAAGTNARLGKGEYIVAEADESDASFLHLTPIMSVVTNIDEDHMDTYGHSVEKLHQAFID
                        190       200       210       220       230       240

                        250       260       270       280       290       300
m092.pep         FIHRMPFYGKAFLCIDSEHVRAILPKVSKPYATYGLDDTADIYATDIENVGAQMKFTVHV
                 ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
g092             FIHRMPFYGKAFLCVDSEHVRAILPKVSKPYATYGLDDTADIYATDIENVGAQMKFTVHV
                        250       260       270       280       290       300

                        310       320       330       340       350       360
m092.pep         QMKGHEQGSFEVVLNMPGRHNVLNALAAIGVALEVGASVEAIQKGLLGFEGVGRRFQKYG
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g092             QMKGHEQGSFEVVLNMPGRHNVLNALAAIGVALEVGASVEAIQKGLLGFEGVGRRFQKYG
                        310       320       330       340       350       360

                        370       380       390       400       410       420
m092.pep         DIKLPNGGTALLVDDYGHHPVEMAATLAAARGAYLEKRLVLAFQPHRYTRTRDLFEDFTK
                 |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
g092             DIKLPNGGTALLVDDYGHHPVEMAATLAAARGAYPEKRLVLAFQPHRYTRTRDLFEDFTK
                        370       380       390       400       410       420

                        430       440       450       460       470       480
m092.pep         VLNTVDALVLTEVYAAGEEPIAAADSRALARAIRVLGKLEPIYCENVADLPEMLLNVLQD
                 ||||||||||||||||||||:||||||||||||||||||||||||||||||:||:|||||
g092             VLNTVDALVLTEVYAAGEEPVAAADSRALARAIRVLGKLEPIYCENVADLPQMLMNVLQD
                        430       440       450       460       470       480

                        490       500
m092.pep         GDIVLNMGAGSINRVPAALLALSKQIX
                 ||:|||||||||||||:||| |||||
g092             GDVVLNMGAGSINRVPSALLELSKQIX
                        490       500
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 313>:

a092.seq

```
  1  ATGTTTTTTA TTTCAATCCG CTATATATTT GTCAGAAAAC TATGGCGCGC
 51  AAACGGTCAG CCCTTTAAAA TAACGCCTTT ACGCATCGAA AATCCACCGG
101  AACGCAACAT TATGATGAAA AATCGAGTGA CCAACATCCA TTTTGTCGGT
151  ATCGGCGGCG TCGGCATGAG CGGTATCGCC GAAGTCTTGC ACAATTTGGG
201  TTTTAAAGTT TCCGGTTCGG ATCAGGCGCG AAATGCCGCT ACCGAGCATT
251  TGGGCAGCCT GGGCATTCAA GTTTATCCCG GCCATACCGC AGAACACGTT
```

```
301  AACGGTGCGG ATGTCGTCGT TACCTCTACC GCCGTCAAAA AAGAAAATCC
351  CGAAGTTGTC GCTGCGTTGG AGCAGCAAAT TCCCGTTATT CCGCGCGCCC
401  TGATGTTGGC GGAGTTGATG CGCTTCCGTG ACGGCATCGC CATTGCCGGC
451  ACGCACGGCA AAACCACGAC CACCAGCCTG ACCGCCTCCA TCCTCGGCGC
501  GGCAGGACTT GACCCGACTT TCGTTATCGG CGGCAAACTC AACGCCGCAG
551  GCACCAACGC CCGCTTGGGC AAAGGCGAAT ACATCGTTGC CGAAGCCGAC
601  GAGTCGGATG CATCCTTTCT GCACCTGACA CCGATTATGT CCGTCGTTAC
651  CAATATCGAC GAAGACCATA TGGATACCTA CGGGCACAGT GTTGAGAAGC
701  TGCATCAGGC GTTTATCGAT TTCATCCACC GTATGCCCTT CTACGGCAAA
751  GCCTTTTTGT GTATTGACAG CGAACACGTC CGCGCGATTT TGCCCAAAGT
801  GAGCAAACCT TATGCTACTT ACGGTTTGGA CGATACCGCC GACATCTACG
851  CCACCGACAT CGAAAACGTC GGCGCGCAAA TGAAATTCAC CGTCCATGTT
901  CAAATGAAAG GACATGAGCA GGGGTCGTTT GAAGTCGTGC TGAATATGCC
951  CGGCAGACAC AACGTGCTGA ACGCATTGGC AGCCATCGGC GTGGCGCTGG
1001 AAGTCGGCGC ATCGGTTGAA GCGATCCAAA AAGGCTTGCT CGGCTTTGAA
1051 GGTGTCGGCC GCCGCTTCCA AAAATACGGC GACATCAAGT TGCCAAACGG
1101 TGGAACCGCG CTCTTGGTGG ACGACTACGG ACACCACCCC GTCGAAATGG
1151 CGGCGACCCT TTCCGCCGCA CGCGGCGCGT ATCCGGAAAA ACGTTTGGTA
1201 CTCGCCTTCC AGCCGCACCG CTATACCCGC ACGCGCGATT TGTTTGAAGA
1251 CTTTACCAAA GTCCTCAATA CCGTTGACGC GCTGGTGCTG ACCGAAGTTT
1301 ATGCCGCCGG TGAAGAGCCG ATTGCCGCCG CTGATTCCCG CGCTCTTGCC
1351 CGCGCCATCC GCGTGTTGGG CAAACTCGAG CCGATTTACT GCGAAAACGT
1401 TGCCGATCTG CCCGAAATGC TGTTGAACGT TTTGCAGGAC GGCGACATCG
1451 TGTTGAATAT CGGTGCGGGA AGCATCAACC GCGTCCCCGC CGCGCTGCTG
1501 GAATTGTCGA AACAGATTTG A
```

This corresponds to the amino acid sequence <SEQ ID 314; ORF 092.a>:

a092.pep

```
  1  MFFISIRYIF VRKLWRANGQ PFKITPLRIE NPPERNIMMK NRVTNIHFVG
 51  IGGVGMSGIA EVLHNLGFKV SGSDQARNAA TEHLGSLGIQ VYPGHTAEHV
101  NGADVVVTST AVKKENPEVV AALEQQIPVI PRALMLAELM RFRDGIAIAG
151  THGKTTTTSL TASILGAAGL DPTFVIGGKL NAAGTNARLG KGEYIVAEAD
201  ESDASFLHLT PIMSVVTNID EDHMDTYGHS VEKLHQAFID FIHRMPFYGK
251  AFLCIDSEHV RAILPKVSKP YATYGLDDTA DIYATDIENV GAQMKFTVHV
301  QMKGHEQGSF EVVLNMPGRH NVLNALAAIG VALEVGASVE AIQKGLLGFE
351  GVGRRFQKYG DIKLPNGGTA LLVDDYGHHP VEMAATLSAA RGAYPEKRLV
401  LAFQPHRYTR TRDLFEDFTK VLNTVDALVL TEVYAAGEEP IAAADSRALA
451  RAIRVLGKLE PIYCENVADL PEMLLNVLQD GDIVLNMGAG SINRVPAALL
501  ELSKQI*
```

m092/a092 99.4% identity over a 506 aa overlap

```
                 10        20        30        40        50        60
m092.pep   MFFISIRYIFVRKLWRANGQPFKITPLRIENPPERNIMMKNRVTNIHFVGIGGVGMSGIA
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a092       MFFISIRYIFVRKLWRANGQPFKITPLRIENPPERNIMMKNRVTNIHFVGIGGVGMSGIA
                 10        20        30        40        50        60

                 70        80        90       100       110       120
m092.pep   EVLHNLGFKVSGSDQARNAATEHLGSLGIQVYPGHTAEHVNGADVVVTSTAVKKENPEVV
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a092       EVLHNLGFKVSGSDQARNAATEHLGSLGIQVYPGHTAEHVNGADVVVTSTAVKKENPEVV
                 70        80        90       100       110       120

                130       140       150       160       170       180
m092.pep   AALEQQIPVIPRALMLAELMRFRDGIAIAGTHGKTTTTSLTASILGAAGLDPTFVIGGKL
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a092       AALEQQIPVIPRALMLAELMRFRDGIAIAGTHGKTTTTSLTASILGAAGLDPTFVIGGKL
                130       140       150       160       170       180

                190       200       210       220       230       240
m092.pep   NAAGTNARLGKGEYIVAEADESDASFLHLTPIMSVVTNIDEDHMDTYGHSVEKLHQAFID
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a092       NAAGTNARLGKGEYIVAEADESDASFLHLTPIMSVVTNIDEDHMDTYGHSVEKLHQAFID
```

```
                  190       200       210       220       230       240

                  250       260       270       280       290       300
m092.pep   FIHRMPFYGKAFLCIDSEHVRAILPKVSKPYATYGLDDTADIYATDIENVGAQMKFTVHV
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a092       FIHRMPFYGKAFLCIDSEHVRAILPKVSKPYATYGLDDTADIYATDIENVGAQMKFTVHV
                  250       260       270       280       290       300

                  310       320       330       340       350       360
m092.pep   QMKGHEQGSFEVVLNMPGRHNVLNALAAIGVALEVGASVEAIQKGLLGFEGVGRRFQKYG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a092       QMKGHEQGSFEVVLNMPGRHNVLNALAAIGVALEVGASVEAIQKGLLGFEGVGRRFQKYG
                  310       320       330       340       350       360

                  370       380       390       400       410       420
m092.pep   DIKLPNGGTALLVDDYGHHPVEMAATLAAARGAYLEKRLVLAFQPHRYTRTRDLFEDFTK
           |||||||||||||||||||||||||||:|||||| ||||||||||||||||||||||||
a092       DIKLPNGGTALLVDDYGHHPVEMAATLSAARGAYPEKRLVLAFQPHRYTRTRDLFEDFTK
                  370       380       390       400       410       420

                  430       440       450       460       470       480
m092.pep   VLNTVDALVLTEVYAAGEEPIAAADSRALARAIRVLGKLEPIYCENVADLPEMLLNVLQD
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a092       VLNTVDALVLTEVYAAGEEPIAAADSRALARAIRVLGKLEPIYCENVADLPEMLLNVLQD
                  430       440       450       460       470       480

                  490       500
m092.pep   GDIVLNMGAGSINRVPAALLALSKQIX
           |||||||||||||||||||| ||||||
a092       GDIVLNMGAGSINRVPAALLELSKQIX
                  490       500
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 315>:

```
g093.seq
      1  aTGCAGAATt ttgGCAAAGT ggccgtATTG ATGGGtggtT TTTCCAGCGA
     51  ACGAGAaatc tcgcTGGACA GCgGTACCGC CATTTTGAAC GCCTTAAAAA
    101  GCAAAGGCAT AGACGCATAC GCCTTCGACC CTAAGGAAAC GCCGTTATCC
    151  GAACTGAAGG AGCGGGGCTT TCAGACGGCA TTCAACATCC TTCACGGTAC
    201  TTACGGCGAA GACGGGGCTG TTCAGGGTGC ATTGGAACTG TTGGGCATTC
    251  CCTATACCGG CAGCGGTGTC GCCGCCTCCG CCATCGGCAT GGACAAATAC
    301  CGCTGCAAAC TGATTTGGCA GGCATTGGGA TTACCCGTTC CCGAGTTCGC
    351  CGTACTGTAC GATGATACCG ATTTCGATGC CGTCGAAGAA AAATTGGGTC
    401  TGCCGATGTT TGTGAAGCCG GCGGCCGAAG GCAGCAGCgt cggcgtggta
    451  aAAGTCAAAG AAAaaggccg TCTGAAAAGC GTTtacgaag aatTGAaaCA
    501  CCTTcagggg cgaAAtcatt gccgAacgTT TTATCGGCGG CGGCGAATAT
    551  TCCTGCCCCG TCCTGAACGG CAAAGGGCTG CCCGGCATAC ACATCATCCC
    601  CGCAACCGAG TTTTACGAct acgaagccaa GtacaaCCGA GACGAcacca
    651  tttaTCAATG TCCTTCGGAA GATTTGACCG AAGCCGAAGA AAGCCTGATG
    701  CGCGAACTGG CGGTTCGCGG CGCACAGGCA ATCGGTGCGG AAGGCTGCGT
    751  GCGCGTCGAT TTCCTCAAAG ATACCGACGG CAAACTCTAT CTGTTGGAAA
    801  TCAACACCCT GCCCGGTATG ACCGGCCATA G
```

This corresponds to the amino acid sequence <SEQ ID 316; ORF 093.ng>:

```
g093.pep
      1  MQNFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPLS
     51  ELKERGFQTA FNILHGTYGE DGAVQGALEL LGIPYTGSGV AASAIGMDKY
    101  RCKLIWQALG LPVPEFAVLY DDTDFDAVEE KLGLPMFVKP AAEGSSVGVV
    151  KVKEKGRLKS VYEELKHLQG RNHCRTFYRR RRIFLPRPER QRAARHTHHP
    201  RNRVLRLRSQ VQPRRHHLSM SFGRFDRSRR KPDARTGGSR RTGNRCGRLR
    251  ARRFPQRYRR QTLSVGNQHP ARYDRP*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 317>:

```
m093.seq
```

```
  1  ATGCAGAATT TTGGCAAAGT GGCCGTATTG ATGGGCGGTT TTTCCAGCGA
 51  ACGAGAAATC TCGCTGGACA GCGGCACCGC CATTTTGAAT GCTTTAAAAA
101  GCAAAGGCAT AGACGCATAC GCCTTCGATC CTAAAGAAAC CCCATTGTCT
151  GAATTGAAGG CACAAGGTTT TCAGACGGCA TTCAACATCC TTCACGGTAC
201  TTACGGCrAA GACGGGGCGG TTCAGGGTGC ATTGGAACTG TTGGGCATTC
251  CCTATACCGG CAGCGGTGTC GCCGCATCCG CCATCGGCAT GGACAAATAC
301  CGCTGCAAAC TGATTTGGCA GGCATTGGGA TTGCCCGTTC CCGAGTTCGC
351  CGTCCTGCAC GACGACACTG ATTTCGATGC CGTCGAAGAA AAATTGGGCC
401  TGCCGATGTT TGTGAAACCG GCGGCCGAAG GCAGCAGCGT AGGCGTGGTA
451  AAAGTCAAAG GAAAAGGCCG TCTGAAAAGC GTTTACGAAG AATTGAAACA
501  CCTTCAGGG. CGAAATCATT GCCGAACGTT TTATCGGCGG CGGCGAATAT
551  TCCTGCCCCG TCCTGAACGG CAAAGGGCTG CCCGGCATAC ACATCATTCC
601  CGCAACCGAG TTTTACGACT ACGAAGCCAA GTACAACCGC GACGACACCA
651  TTTATCAATG TCCTTCGGAA GATTTGACCG AAGCCGAAGA AAGCCTGATG
701  CGCGAACTGG CGGTTCGCGG CGCGCAGGCA ATCGGTGCGG AAGGCTGCGT
751  GCGCGTCGAT TTCCTCAAAG ATACCGACGG CAAACTCTAT CTGTTGGAAA
801  TCAACACCCT GCCCGGTATG ACGAGCCATA G
```

This corresponds to the amino acid sequence <SEQ ID 318; ORF 093>:

m093.pep

```
  1  MQNFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPLS
 51  ELKAQGFQTA FNILHGTYGX DGAVQGALEL LGIPYTGSGV AASAIGMDKY
101  RCKLIWQALG LPVPEFAVLH DDTDFDAVEE KLGLPMFVKP AAEGSSVGVV
151  KVKGKGRLKS VYEELKHLQX RNHCRTFYRR RRIFLPRPER QRAARHTHHS
201  RNRVLRLRSQ VQPRRHHLSM SFGRFDRSRR KPDARTGGSR RAGNRCGRLR
251  ARRFPQRYRR QTLSVGNQHP ARYDEP*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 093 shows 96.7% identity over a 276 aa overlap with a predicted ORF (ORF 093.ng) from *N. gonorrhoeae:*

m093/g093

```
                    10        20        30        40        50        60
m093.pep    MQNFGKVAVLMGGFSSEREISLDSGTAILNALKSKGIDAYAFDPKETPLSELKAQGFQTA
            |||||||||||||||||||||||||||||||||||||||||||||||||||||| :||||
g093        MQNFGKVAVLMGGFSSEREISLDSGTAILNALKSKGIDAYAFDPKETPLSELKERGFQTA
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m093.pep    FNILHGTYGXDGAVQGALELLGIPYTGSGVAASAIGMDKYRCKLIWQALGLPVPEFAVLH
            ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||:
g093        FNILHGTYGEDGAVQGALELLGIPYTGSGVAASAIGMDKYRCKLIWQALGLPVPEFAVLY
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m093.pep    DDTDFDAVEEKLGLPMFVKPAAEGSSVGVVKVKGKGRLKSVYEELKHLQXRNHCRTFYRR
            |||||||||||||||||||||||||||||||||| ||||||||||||||| |||||||||
g093        DDTDFDAVEEKLGLPMFVKPAAEGSSVGVVKVKEKGRLKSVYEELKHLQGRNHCRTFYRR
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m093.pep    RRIFLPRPERQRAARHTHHSRNRVLRLRSQVQPRRHHLSMSFGRFDRSRRKPDARTGGSR
            ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
g093        RRIFLPRPERQRAARHTHHPRNRVLRLRSQVQPRRHHLSMSFGRFDRSRRKPDARTGGSR
                   190       200       210       220       230       240

                   250       260       270
m093.pep    RAGNRCGRLRARRFPQRYRRQTLSVGNQHPARYDEPX
            |:||||||||||||||||||||||||||||||||:||
g093        RTGNRCGRLRARRFPQRYRRQTLSVGNQHPARYDRPX
                   250       260       270
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 319>:

a093.seq

```
  1  ATGCAGAATT TTGGCAAAGT GGCCGTATTG ATGGGCGGTT TTTCCAGCGA
 51  ACGAGAAATC TCGCTGGACA GCGGCACCGC CATTTTGAAT GCTTTAAAAA
101  GCAAAGGCAT AGACGCATAC GCCTTCGATC CCAAGGAAAC CCCATTGTCT
151  GAATTGAAGG CACAAGGTTT TCAGACGGCA TTCAACATCC TTCACGGTAC
201  TTACGGCGAA GACGGGGCTG TTCAGGGTGC ATTGGAACTG TTGGGCATTC
251  CCTATACCGG CAGCGGTGTC GCCGCATCCG CCATCGGCAT GGACAAATAC
301  CGCTGCAAAC TGATTTGGCA GGCATTGGGA TTGCCCGTTC CCGAGTTCGC
351  CGTCCTGCAC GACGACACTG ATTTCGATGC CGTCGAAGAA AAATTGGGCC
401  TGCCGATGTT TGTGAAACCG GCGGCCGAAG GCAGCAGCGT AGGCGTGGTA
451  AAAGTCAAAG GAAAAGGCCG TCTGAAAAGC GTTTACGAAG AATTGAAACA
501  CTTTCAGGG. CGAAATCATT GCCGAACGGT TTATCGGCGG CGGCGAATAT
551  TCCTGCCCTG TGTTGAACGG CAAAGGCCTG CCCGGCATAC ACATCATCCC
601  CGCGACCGAG TTTTATGACT ACGAAGCCAA GTACAACCGC AACGACACCA
651  TTTATCAATG TCCTTCGGAA GATCTGACCG AAGCCGAAGA AAGCCTGATG
701  CGCGAACTGG CGGTTCGCGG CGCGCAGGCA ATCGGTGCGG AAGGCTGCGT
751  GCGCGTCGAT TTCCTCAAAG ATACCGACGG CAAACTCTAT CTGTTGGAAA
801  TCAACACCCT GCCCGGTATG ACCGGCCATA G
```

This corresponds to the amino acid sequence <SEQ ID 320; ORF 093.a>:

a093.pep

```
  1  MQNFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPLS
 51  ELKAQGFQTA FNILHGTYGE DGAVQGALEL LGIPYTGSGV AASAIGMDKY
101  RCKLIWQALG LPVPEFAVLH DDTDFDAVEE KLGLPMFVKP AAEGSSVGVV
151  KVKGKGRLKS VYEELKHFQX RNHCRTVYRR RRIFLPCVER QRPARHTHHP
201  RDRVL*LRSQ VQPQRHHLSM SFGRSDRSRR KPDARTGGSR RAGNRCGRLR
251  ARRFPQRYRR QTLSVGNQHP ARYDRP*
```

m093/a093 95.7% identity over a 276 aa overlap

```
                    10        20        30        40        50        60
m093.pep    MQNFGKVAVLMGGFSSEREISLDSGTAILNALKSKGIDAYAFDPKETPLSELKAQGFQTA
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a093        MQNFGKVAVLMGGFSSEREISLDSGTAILNALKSKGIDAYAFDPKETPLSELKAQGFQTA
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m093.pep    FNILHGTYGXDGAVQGALELLGIPYTGSGVAASAIGMDKYRCKLIWQALGLPVPEFAVLH
            ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
a093        FNILHGTYGEDGAVQGALELLGIPYTGSGVAASAIGMDKYRCKLIWQALGLPVPEFAVLH
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m093.pep    DDTDFDAVEEKLGLPMFVKPAAEGSSVGVVKVKGKGRLKSVYEELKHLQXRNHCRTFYRR
            |||||||||||||||||||||||||||||||||||||||||||||||:|||||||| |||
a093        DDTDFDAVEEKLGLPMFVKPAAEGSSVGVVKVKGKGRLKSVYEELKHFQXRNHCRTVYRR
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m093.pep    RRIFLPRPERQRAARHTHHSRNRVLRLRSQVQPRRHHLSMSFGRFDRSRRKPDARTGGSR
            ||||||  |||| |||||| |:||| |||||||:|||||||||| |||||||||||||||
a093        RRIFLPCVERQRPARHTHHPRDRVLXLRSQVQPQRHHLSMSFGRSDRSRRKPDARTGGSR
                   190       200       210       220       230       240

                   250       260       270
m093.pep    RAGNRCGRLRARRFPQRYRRQTLSVGNQHPARYDEPX
            ||||||||||||||||||||||||||||||||||:||
a093        RAGNRCGRLRARRFPQRYRRQTLSVGNQHPARYDRPX
                   250       260       270
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 321>:

```
g094.seq
      1  ATGTATTCGC CTTTGCCCAA GCGGGCGTTG GTGCCTGCGG CGTTGAGTTT
     51  GCCGCCGATA ACGAAAGTGG GGTCGAGTCC TGCCGCGCCG AGGATGGAGG
    101  CGGTCAGGCT GGTGGTCGTG GTTTTGCCGT GCGTAccggc aatggcgatg
    151  cCGTCACGGA AGCGCATCAG CTCTGCCAGC ATCAAGGCGC GCGGAATAAC
    201  GGGAATTTGC CGCTCCAACG CAgcgacaAC TTCGGgattT TCTTTCTTGA
    251  CGGCGGTAGA GGCAACGACG ACATccgcAC CGTTAACGTG TTCTGCGGTA
    301  TGGCCGGGAT AA
```

This corresponds to the amino acid sequence <SEQ ID 322; ORF 094.ng>:

```
g094.pep
      1  MYSPLPKRAL VPAALSLPPI TKVGSSPAAP RMEAVRLVVV VLPCVPAMAM
     51  PSRKRISSAS IKARGITGIC RSNAATTSGF SFLTAVEATT TSAPLTCSAV
    101  WPG*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 323>:

```
m094.seq
      1  ATGTATTCGC CTTTGCCCAA GCGGGCGTTA GTGCCTGCGG CGTTGAGTTT
     51  GCCGCCGATA ACGAAAGTCG GGTCAAGTCC TGCCGCGCCG AGGATGGAGG
    101  CGGTCAGGCT GGTGGTCGTG GTTTTGCCGT GCGTGCCGGC AATGGCGATG
    151  CCGTCACGGA AGCGCATCAA CTCCGCCAAC ATCAGGGCGC GCGGAATAAC
    201  GGGAATTTGC TGCTCCAACG CAGCGACAAC TTCGGGATTT TCTTTTTTGA
    251  CGGCGGTAGA GGTAACGACG ACATCCGCAC CGTTAACGTG TTCGGCGGTA
    301  TGGCCGGGAT AA
```

This corresponds to the amino acid sequence <SEQ ID 324; ORF 094>:

```
m094.pep
      1  MYSPLPKRAL VPAALSLPPI TKVGSSPAAP RMEAVRLVVV VLPCVPAMAM
     51  PSRKRINSAN IRARGITGIC CSNAATTSGF SFLTAVEVTT TSAPLTCSAV
    101  WPG*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 094 shows 95.1% identity over a 103 aa overlap with a predicted ORF (ORF 094.ng) from *N. gonorrhoeae:*

```
m094/g094

                         10        20        30        40        50        60
m094.pep        MYSPLPKRALVPAALSLPPITKVGSSPAAPRMEAVRLVVVVLPCVPAMAMPSRKRINSAN
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||:
g094            MYSPLPKRALVPAALSLPPITKVGSSPAAPRMEAVRLVVVVLPCVPAMAMPSRKRISSAS
                         10        20        30        40        50        60

                         70        80        90       100
m094.pep        IRARGITGICCSNAATTSGFSFLTAVEVTTTSAPLTCSAVWPGX
                |:|||||||| ||||||||||||||||:||||||||||||||||
g094            IKARGITGICRSNAATTSGFSFLTAVEATTTSAPLTCSAVWPGX
                         70        80        90       100
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 325>:

```
a094.seq
      1  ATGTATTCGC CTTTGCCCAA GCGGGCGTTG GTGCCTGCGG CGTTGAGTTT
     51  GCCGCCGATA ACGAAAGTCG GGTCAAGTCC TGCCGCCCCG AGGATGGAGG
    101  CGGTCAGGCT GGTGGTCGTG GTTTTGCCGT GCGTGCCGGC AATGGCGATG
    151  CCGTCACGGA AGCGCATCAA CTCCGCCAAC ATCAGGGCGC GCGGAATAAC
    201  GGGAATTTGC TGCTCCAACG CAGCGACAAC TTCGGGATTT TCTTTTTTGA
    251  CGGCGGTAGA GGTAACGACG ACATCCGCAC CGTTAACGTG TTCTGCGGTA
    301  TGGCCGGGAT AA
```

This corresponds to the amino acid sequence <SEQ ID 326; ORF 094.a>:

a094.pep

```
  1  MYSPLPKRAL VPAALSLPPI TKVGSSPAAP RMEAVRLVVV VLPCVPAMAM
 51  PSRKRINSAN IRARGITGIC CSNAATTSGF SFLTAVEVTT TSAPLTCSAV
101  WPG*
```

m094/a094 100.0% identity over a 103 aa overlap

```
                    10        20        30        40        50        60
m094.pep    MYSPLPKRALVPAALSLPPITKVGSSPAAPRMEAVRLVVVVLPCVPAMAMPSRKRINSAN
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a094        MYSPLPKRALVPAALSLPPITKVGSSPAAPRMEAVRLVVVVLPCVPAMAMPSRKRINSAN
                    10        20        30        40        50        60

                    70        80        90       100
m094.pep    IRARGITGICCSNAATTSGFSFLTAVEVTTTSAPLTCSAVWPGX
            ||||||||||||||||||||||||||||||||||||||||||||
a094        IRARGITGICCSNAATTSGFSFLTAVEVTTTSAPLTCSAVWPGX
                    70        80        90       100
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 327>:

g095.seq

```
  1  ATGTCCTTTC ATTTGAACAT GGACGGTGAA TTTCATTTGC GCGCCGACGT
 51  TTTCGATGTC GGTGGCGTAG ATGTCGGCGG TATCGTCCAA ACCGTAAGTA
101  GCATAAGGTT TGCTCACTTT GGGCAAAATC GCGCGGACGT GTTCGCTGTC
151  AACACACAAA AAGGCTTTGC CGTAGAAGGG CATACGGTGG ATGAAATCGA
201  TAAACGCCTG ATGCAGTTTT TCGACGCTGT GCCCGTAGGT ATCCATATGG
251  TCTTCGTCGA TATTGGTAAC GACGGACATA ATCGGGGTCA GTGTAGGAAA
301  GAGGCATCGG ATCGTCGGCT TCGGCAACGA TGTATTCGCC TTTGCCCAAG
351  CGGGCGTTGG TGCCTGCGGC GTTGA
```

This corresponds to the amino acid sequence <SEQ ID 328; ORF 095.ng>:

g095.pep

```
  1  MSFHLNMDGE FHLRADVFDV GGVDVGGIVQ TVSSIRFAHF GQNRADVFAV
 51  NTQKGFAVEG HTVDEIDKRL MQFFDAVPVG IHMVFVDIGN DGHNRGQCRK
101  EASDRRLRQR CIRLCPSGRW CLRR*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 329>:

m095.seq

```
  1  ATGTCCTTTC ATTTGAACAT GGACGGTGAA TTTCATTTGC GCGCCGACGT
 51  TTTCGATGTC GGTGGCGTAG ATGTCGGCGG TATCGTCCAA ACCGTAAGTA
101  GCATAAGGTT TGCTCACTTT GGGCAAAATC GCGCGGACGT GTTCGCTGTC
151  AATACACAAA AAGGCTTTGC CGTAGAAGGG CATACGGTGG ATGAAATCGA
201  TAAACGCCTG ATGCAGTTTT TCGACGCTGT GCCCGTAGGT ATCCATATGG
251  TCTTCGTCGA TATTGGTAAC GACGGACATA ATCGGTGTCA GTGCAGAAAG
301  GATGCATCCG ACCGTCGGCT TCGGCAACGA TGTATTCGCC TTTGCCCAAG
351  CGGGCGTTAG TGCCTGCGGC GTTGA
```

This corresponds to the amino acid sequence <SEQ ID 330; ORF 095>:

m095.pep

```
  1  MSFHLNMDGE FHLRADVFDV GGVDVGGIVQ TVSSIRFAHF GQNRADVFAV
 51  NTQKGFAVEG HTVDEIDKRL MQFFDAVPVG IHMVFVDIGN DGHNRCQCRK
101  DASDRRLRQR CIRLCPSGRX CLRR*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 095 shows 97.6% identity over a 124 aa overlap with a predicted ORF (ORF 095.ng) from *N. gonorrhoeae:*

m095/g095

```
                    10        20        30        40        50        60
m095.pep    MSFHLNMDGEFHLRADVFDVGGVDVGGIVQTVSSIRFAHFGQNRADVFAVNTQKGFAVEG
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
g095          MSFHLNMDGEFHLRADVFDVGGVDVGGIVQTVSSIRFAHFGQNRADVFAVNTQKGFAVEG
                        10        20        30        40        50        60

                        70        80        90       100       110       120
m095.pep      HTVDEIDKRLMQFFDAVPVGIHMVFVDIGNDGHNRCQCRKDASDRRLRQRCIRLCPSGRX
              ||||||||||||||||||||||||||||||||||| ||||:||||||||||||||||||
g095          HTVDEIDKRLMQFFDAVPVGIHMVFVDIGNDGHNRGQCRKEASDRRLRQRCIRLCPSGRW
                        70        80        90       100       110       120

m095.pep      CLRRX
              |||||
g095          CLRRX
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 331>:

```
a095.seq
      1  ATGTCCTTTC ATTTGAACAT GGACGGTGAA TTTCATTTGC GCGCCGACGT
     51  TTTCGATGTC GGTGGCGTAG ATGTCGGCGG TATCGTCCAA ACCGTAAGTA
    101  GCATAAGGTT TGCTCACTTT GGGCAAAATC GCGCGGACGT GTTCGCTGTC
    151  AATACACAAA AAGGCTTTGC CGTAGAAGGG CATACGGTGG ATGAAATCGA
    201  TAAACGCCTG ATGCAGCTTC TCAACACTGT GCCCGTAGGT ATCCATATGG
    251  TCTTCGTCGA TATTGGTAAC GACGGACATA ATCGGTGTCA GTGCAGAAAG
    301  GATGCATCCG ACCGTCGGCT TCGGCAACGA TGTATTCGCC TTTGCCCAAG
    351  CGGGCGTTGG TGCCTGCGGC GTTGA
```

This corresponds to the amino acid sequence <SEQ ID 332; ORF 095.a>:

```
a095.pep
      1  MSFHLNMDGE FHLRADVFDV GGVDVGGIVQ TVSSIRFAHF GQNRADVFAV
     51  NTQKGFAVEG HTVDEIDKRL MQLLNTVPVG IHMVFVDIGN DGHNRCQCRK
    101  DASDRRLRQR CIRLCPSGRW CLRR*

m095/a095     96.0% identity in 124 aa overlap

                        10        20        30        40        50        60
m095.pep      MSFHLNMDGEFHLRADVFDVGGVDVGGIVQTVSSIRFAHFGQNRADVFAVNTQKGFAVEG
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a095          MSFHLNMDGEFHLRADVFDVGGVDVGGIVQTVSSIRFAHFGQNRADVFAVNTQKGFAVEG
                        10        20        30        40        50        60

                        70        80        90       100       110       120
m095.pep      HTVDEIDKRLMQFFDAVPVGIHMVFVDIGNDGHNRCQCRKDASDRRLRQRCIRLCPSGRX
              ||||||||||||::::|||||||||||||||||||||||||||||||||||||||||||
a095          HTVDEIDKRLMQLLNTVPVGIHMVFVDIGNDGHNRCQCRKDASDRRLRQRCIRLCPSGRW
                        70        80        90       100       110       120

m095.pep      CLRRX
              |||||
a095          CLRRX
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 333>:

```
g096.seq
      1  ATGGCCGGTC ATACCGGGCA GGGTGTTGAT TTCCAACAGA TAGAGTTTGC
     51  CGTCGGTATC TTTGAGGAAA TCGACGCGCA CGCAGCCTTC CGCACCGATT
    101  GCCTGTGCGC CGCGAACCGC CAGTTCGCGC ATCAGGCTTT CTTCGGCTTC
    151  GGTCAAATCT TCCGAAGGAC ATTGAtaaat ggtgTCGTCT CGGttgtaCt
```

```
201   tggcttcgta gTCGTAAAAC TCGGTTGCGG GGATGATGTG TATGCCGGGC
251   AGCCCTTTGC CGTTCAGGAC GGGGCAGGAA TATTCGCCGC CGCCGATAAA
301   AcgtTcggca atgaTTtcgc ccctgAAGGT GttTCAattc ttcgtaAACG
351   CTTTTCAGAc ggcctTTTC TTTGA
```

This corresponds to the amino acid sequence <SEQ ID 334; ORF 096.ng>:

g096.pep

```
  1   MAGHTGQGVD FQQIEFAVGI FEEIDAHAAF RTDCLCAANR QFAHQAFFGF
 51   GQIFRRTLIN GVVSVVLGFV VVKLGCGDDV YAGQPFAVQD GAGIFAAADK
101   TFGNDFAPEG VSILRKRFSD GLFL*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 335>:

m096.seq

```
  1   ATGGCTCGTC ATACCGGGCA GGGTGTTGAT TTCCAACAGA TAGAGTTTGC
 51   CGTCGGTATC TTTGAGGAAA TCGACGCGCA CGCAGCCTTC CGCACCGATT
101   GCCTGCGCGC CGCGAACCGC CAGTTCGCGC ATCAGGCTTT CTTCGGCTTC
151   GGTCAAATCT TCCGAAGGAC ATTGATAAAT GGTGTCGTCG CGGTTGTACT
201   TGGCTTCGTA GTCGTAAAAC TCGGTTGCGG GAATGATGTG TATGCCGGGC
251   AGCCCTTTGC CGTTCAGGAC GGGGCAGGAA TATTCGCCGC CGCCGATAAA
301   ACGTTCGGCA ATGATTTCGC CC.TGAAGGT GTTTCAATTC TTCGTAAACG
351   CTTTTCAGAC GGCCTTTTCC TTTGA
```

This corresponds to the amino acid sequence <SEQ ID 336; ORF 096>:

m096.pep

```
  1   MARHTGQGVD FQQIEFAVGI FEEIDAHAAF RTDCLRAANR QFAHQAFFGF
 51   GQIFRRTLIN GVVAVVLGFV VVKLGCGNDV YAGQPFAVQD GAGIFAAADK
101   TFGNDFAXEG VSILRKRFSD GLFL*
```

m096/g096   96.0% identity in 124 aa overlap

```
                       10        20        30        40        50        60
m096.pep       MARHTGQGVDFQQIEFAVGIFEEIDAHAAFRTDCLRAANRQFAHQAFFGFGQIFRRTLIN
               || |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
g096           MAGHTGQGVDFQQIEFAVGIFEEIDAHAAFRTDCLCAANRQFAHQAFFGFGQIFRRTLIN
                       10        20        30        40        50        60

                       70        80        90       100       110       120
m096.pep       GVVAVVLGFVVVKLGCGNDVYAGQPFAVQDGAGIFAAADKTFGNDFAXEGVSILRKRFSD
               |||:|||||||||||||:|||||||||||||||||||||||||||||| |||||||||||
g096           GVVSVVLGFVVVKLGCGDDVYAGQPFAVQDGAGIFAAADKTFGNDFAPEGVSILRKRFSD
                       70        80        90       100       110       120

m096.pep       GLFLX
               |||||
g096           GLFLX
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 337>:

a096.seq

```
  1   ATGGCCGGTC ATACCGGGCA GGGTGTTGAT TTCCAACAGA TAGAGTTTGC
 51   CGTCGGTATC TTTGAGGAAA TCGACGCGCA CGCAGCCTTC CGCACCGATT
101   GCCTGCGCGC CGCGAACCGC CAGTTCGCGC ATCAGGCTTT CTTCGGCTTC
151   GGTCAGATCT TCCGAAGGAC ATTGATAAAT GGTGTCGTTG CGGTTGTACT
201   TGGCTTCGTA GTCATAAAAC TCGGTCGCGG GGATGATGTG TATGCCGGGC
251   AGGCCTTTGC CGTTCAACAC AGGGCAGGAA TATTCGCCGC CGCCGATAAA
301   CCGTTCGGCA ATGATTTCGC CCT.GAAAGT GTTTCAATTC TTCGTAAACG
351   CTTTTCAGAC GGCCTTTTCC TTTGA
```

This corresponds to the amino acid sequence <SEQ ID 338; ORF 096.ng>:

a096.pep

```
  1   MAGHTGQGVD FQQIEFAVGI FEEIDAHAAF RTDCLRAANR QFAHQAFFGF
```

```
 51  GQIFRRTLIN GVVAVVLGFV VIKLGRGDDV YAGQAFAVQH RAGIFAAADK
101  PFGNDFAXES VSILRKRFSD GLFL*
```

m096/a096   92.7% identity in 124 aa overlap

```
                    10        20        30        40        50        60
m096.pep    MARHTGQGVDFQQIEFAVGIFEEIDAHAAFRTDCLRAANRQFAHQAFFGFGQIFRRTLIN
            || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a096        MAGHTGQGVDFQQIEFAVGIFEEIDAHAAFRTDCLRAANRQFAHQAFFGFGQIFRRTLIN
                    10        20        30        40        50        60

                    70        80        9.       100       110       120
m096.pep    GVVAVVLGFVVVKLGCGNDVYAGQPFAVQDGAGIFAAADKTFGNDFAXEGVSILRKRFSD
            ||||||||||:||| |:||||| ||||  |||||||| ||||||||:|||||||||
a096        GVVAVVLGFVVIKLGRGDDVYAGQAFAVQHRAGIFAAADKPFGNDFAXESVSILRKRFSD
                    70        80        90       100       110       120

m096.pep    GLFLX
            |||||
a096        GLFLX
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 339>:

```
g097.seq
   1  ATGGATATTT CAAAACAAAC ATTGCTGGAT AGGGTTTTTA ACCTGAAGGC
  51  AAACGGTACG ACGGTACGTA CCGAGTTGAT GGCGGGTTTG ACGACCTTTT
 101  TGACGATGTG CTACATCGTT ATCGTCAATC CCCTGATTTT GGGCGAGACC
 151  GGAATGGATA TGGGGGCGGT ATTCGTCGCT ACCTGTATCG CATCCGCCAT
 201  CGGCTGTTTT GTCATGGGTT TTATCGGCAA CTATCCGATT GCGCTTGCCC
 251  CGGGGATGGG GCTGAATGCC TATTTCACCT TTGCCGTCGT TAAGGGTATG
 301  GGCGTGCCTT GGCAGGTGGC GTTGGGTGCG GTGTTCATTT CCGGTCTGAT
 351  TTTCATCCTG TTCAGCTTTT TTAAAGTCAG GGAAATGCTG GTCAACGCAC
 401  TGCCTATGGG TTTGAAAATG TCGATTGCCG CCGGTATCGG TTTGTTTTTG
 451  GCACTGATTT CCCTGAAAGG CGCAGGCATT ATCGTTGCCA ATCCGGCAAC
 501  CTTGGTCGGC TTGGGCGATA TTCATCAGCC CAGCGCACTG TTGGCATTGT
 551  TCGGTTTTGT CATGGTGGTC GTATTGGGGT ATTTCCGCGT TCAAGGCGCA
 601  ATCATCATCA CCATTCTGAC GATTACCGTC ATTGCCAGCC TGATGGGTTT
 651  GAACGAGTTT CACGGCGTGG TCGGCGAAGT ACCGGGCATT GCGCCGACCT
 701  TTATGCAGAT GGATTTTAAA GGTCTGTTTA CCGTCAGCAT GGTCAGCGTG
 751  ATTTTCGTCT TCTTCTTGGT CGATTTGTTC GACAGTACCG GAACGCTGGT
 801  CGGCGTATCC CACCGTGCCG GACTGCTGGT GGACGGTAAG CTGCCCCGCC
 851  TGAAACGCGC ACTGCTTGCA GACTCTACCG CCATTGTGGC AGGTGCGGCT
 901  TTGGGTACTT CTTCAACCAC GCCTTATGTG GAAAGCGCGG CGGGCGTATC
 951  GGCAGGCGGA CGGACCGGCC TGACGGCGGT TACCGTCGGC GTATTGATGC
1001  TGGCGTGTCT GATGTTCTCC CCATTGGCGA AAAGTGTTCC GGTATTTGCC
1051  ACCGCGCCCG CACTGCTTTA TGTCGGCACG CAGATGCTCC GCAGTGCGAG
1101  GGACATTGAT TGGGACGATA TGACTGAAGC CGCGCCCGCG TTCCTGACCA
1151  TTGTCTTCAT GCCGTTTACC TATTCGATTG CAGACGGCAT CGCCTTCGGC
1201  TTCATCAGCT ATGCCGTGGT CAAACTTTTG TGTCGCCGGA CTGGGGACGT
1251  GCCGCCTATG GTATGGGTTG TTGCCGTATT GTGGGCATTG AAATTCTGGT
1301  ATTTGGGCTG A
```

This corresponds to the amino acid sequence <SEQ ID 340; ORF 097.ng>:

```
g097.pep
   1  MDISKQTLLD RVFNLKANGT TVRTELMAGL TTFLTMCYIV IVNPLILGET
  51  GMDMGAVFVA TCIASAIGCF VMGFIGNYPI ALAPGMGLNA YFTFAVVKGM
 101  GVPWQVALGA VFISGLIFIL FSFFKVREML VNALPMGLKM SIAAGIGLFL
 151  ALISLKGAGI IVANPATLVG LGDIHQPSAL LALFGFVMVV VLGYFRVQGA
 201  IIITILTITV IASLMGLNEF HGVVGEVPGI APTFMQMDFK GLFTVSMVSV
 251  IFVFFLVDLF DSTGTLVGVS HRAGLLVDGK LPRLKRALLA DSTAIVAGAA
 301  LGTSSTTPYV ESAAGVSAGG RTGLTAVTVG VLMLACLMFS PLAKSVPVFA
 351  TAPALLYVGT QMLRSARDID WDDMTEAAPA FLTIVFMPFT YSIADGIAFG
```

```
401  FISYAVVKLL CRRTGDVPPM VWVVAVLWAL KFWYLG*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 341>:

```
m097.seq
   1  ATGGACACTT CAAAACAAAC ACTGTTGGAC GGGATTTTTA AGCTGAAGGC
  51  AAACGGTACk ACGGTGCGTA CCGAGTTGAT GGCGGGTTTG ACAACTTTTT
 101  TGACGATGTG CTACATCGTT ATCGTCAACC CTCyGATTTT GGGCGAGACC
 151  GGCATGGATA TGGGGGCGGT ATTCGTCGCT ACCTGTATCG CGTCTGCCAT
 201  CGGCTGTTTT GTTATGGGTT TTGTCGGCAA CTATCCGATT GCACTCGCAC
 251  CGGGGATGGG GCTGAATGCC TATTTCACCT TTGCCGTCGT TAAGGGTATG
 301  GGCGTGCCTT GGCAGGTTGC GTTGGGTGCG GTGTTCATCT CCGGTCTGAT
 351  TTTTATCCTG TTCAGCTTTT TTAAAGTCAG GGAAATGCTG GTCAACGCAC
 401  TGCCTATGGG TTTGAAAATG TCGATTGCTG CCGGTATCGG TTTGTTTTTG
 451  GCACTGATTT CCCTGAAAGG CGCAGGCATT ATCGTTGCCA ATCCGGCAAC
 501  CTTGGTCGGT TTGGGCGATA TTCATCAGCC GTCCGCGTTG TTGGCATTGT
 551  TCGGTTTTGC TATGGTGGTC GTATTGGGAC ATTTCCGCGT TCAAGGCGCA
 601  ATCATCATCA CCATCTTGAC CATTACCGTC ATTGCCAGCC TGATGGGTTT
 651  GAATGAATTT CACGGCATCA TCGGCGAAGT ACCGAGCATT GCGCCGACTT
 701  TTATGCAGAT GGATTTTGAA GGCCTGTTTA CCGTCAGCAT GGTCAGTGTG
 751  ATTTTCGTCT TCTTCTTGGT CGATCTATTT GACAGTACCG GAACGCTGGT
 801  CGGCATATCC CACCGTGCCG GGCTGCTGGT GGACGGTAAG CTGCCCCGCC
 851  TGAAACGCGC ACTGCTTGCA GACTCTACCG CCATTGTGGC AGGTGCGGCT
 901  TTGGGTACTT CTTCCACCAC GCCTTATGTG GAAAGCGCGG CGGGCGTATC
 951  GGCAGGCGGA CGGACCGGCC TGACGGCGGT TACCGTCGGC GTATTGATGC
1001  TCGCCTGCCT GATGTTTTCA CCTTTGGCGA AAAGTGTTCC CGCTTTTGCC
1051  ACCGCGCCCG CCCTGCTTTA TGTCGGCACG CAGATGCTCC GCAGTGCGAG
1101  GGATATTGAT TGGGACGATA TGACGGAAGC CGCACCTGCG TTCCTGACCA
1151  TTGTTTTCAT GCCGTTTACT TATTCGATTG CAGACGGCAT CGCTTTCGGC
1201  TTCATCAGTT ATGCCGTGGT TAAACTTTTA TGCCGCCGCA CCAAAGACGT
1251  TCCGCCTATG GTATGGATTG TTGCCGTATT GTGGGCACTG AAATTCTGGT
1301  ATTTGGGCTG A
```

This corresponds to the amino acid sequence <SEQ ID 342; ORF 097>:

```
m097.pep
   1  MDTSKQTLLD GIFKLKANGT TVRTELMAGL TTFLTMCYIV IVNPXILGET
  51  GMDMGAVFVA TCIASAIGCF VMGFVGNYPI ALAPGMGLNA YFTFAVVKGM
 101  GVPWQVALGA VFISGLIFIL FSFFKVREML VNALPMGLKM SIAAGIGLFL
 151  ALISLKGAGI IVANPATLVG LGDIHQPSAL LALFGFAMVV VLGHFRVQGA
 201  IIITILTITV IASLMGLNEF HGIIGEVPSI APTFMQMDFE GLFTVSMVSV
 251  IFVFFLVDLF DSTGTLVGIS HRAGLLVDGK LPRLKRALLA DSTAIVAGAA
 301  LGTSSTTPYV ESAAGVSAGG RTGLTAVTVG VLMLACLMFS PLAKSVPAFA
 351  TAPALLYVGT QMLRSARDID WDDMTEAAPA FLTIVFMPFT YSIADGIAFG
 401  FISYAVVKLL CRRTKDVPPM VWIVAVLWAL KFWYLG*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 097 shows 96.3% identity over a 436 aa overlap with a predicted ORF (ORF 097.ng) from *N. gonorrhoeae:*

```
m097/g097

                     10        20        30        40        50        60
m097.pep    MDTSKQTLLDGIFKLKANGTTVRTELMAGLTTFLTMCYIVIVNPXILGETGMDMGAVFVA
            || ||||||| :|:||||||||||||||||||||||||||||| ||||||||||||||||
g097        MDISKQTLLDRVFNLKANGTTVRTELMAGLTTFLTMCYIVIVNPLILGETGMDMGAVFVA
                     10        20        30        40        50        60

                     70        80        90       100       110       120
m097.pep    TCIASAIGCFVMGFVGNYPIALAPGMGLNAYFTFAVVKGMGVPWQVALGAVFISGLIFIL
            ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
g097        TCIASAIGCFVMGFIGNYPIALAPGMGLNAYFTFAVVKGMGVPWQVALGAVFISGLIFIL
                     70        80        90       100       110       120

                    130       140       150       160       170       180
```

```
m097.pep    FSFFKVREMLVNALPMGLKMSIAAGIGLFLALISLKGAGIIVANPATLVGLGDIHQPSAL
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g097        FSFFKVREMLVNALPMGLKMSIAAGIGLFLALISLKGAGIIVANPATLVGLGDIHQPSAL
                    130       140       150       160       170       180

                    190       200       210       220       230       240
m097.pep    LALFGFAMVVVLGHFRVQGAIIITILTITVIASLMGLNEFHGIIGEVPSIAPTFMQMDFE
            ||||||:||||||:||||||||||||||||||||||||||||::|||||:||||||||||:
g097        LALFGFVMVVVLGYFRVQGAIIITILTITVIASLMGLNEFHGVVGEVPGIAPTFMQMDFK
                    190       200       210       220       230       240

                    250       260       270       280       290       300
m097.pep    GLFTVSMVSVIFVFFLVDLFDSTGTLVGISHRAGLLVDGKLPRLKRALLADSTAIVAGAA
            ||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
g097        GLFTVSMVSVIFVFFLVDLFDSTGTLVGVSHRAGLLVDGKLPRLKRALLADSTAIVAGAA
                    250       260       270       280       290       300

                    310       320       330       340       350       360
m097.pep    LGTSSTTPYVESAAGVSAGGRTGLTAVTVGVLMLACLMFSPLAKSVPAFATAPALLYVGT
            ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
g097        LGTSSTTPYVESAAGVSAGGRTGLTAVTVGVLMLACLMFSPLAKSVPVFATAPALLYVGT
                    310       320       330       340       350       360

                    370       380       390       400       410       420
m097.pep    QMLRSARDIDWDDMTEAAPAFLTIVFMPFTYSIADGIAFGFISYAVVKLLCRRTKDVPPM
            |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
g097        QMLRSARDIDWDDMTEAAPAFLTIVFMPFTYSIADGIAFGFISYAVVKLLCRRTGDVPPM
                    370       380       390       400       410       420

                    430
m097.pep    VWIVAVLWALKFWYLGX
            ||:|||||||||||||
g097        VWVVAVLWALKFWYLGX
                    430
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 343>

a097.seq

```
   1  ATGGACACTT CAAAACAAAC ACTGTTGGAC GGGATTTTTA AGCTGAAGGC
  51  AAACGGTACG ACGGTGCGTA CCGAGTTGAT GGCGGGTTTG ACAACTTTTT
 101  TGACGATGTG CTACATCGTT ATCGTCAACC CTCTGATTTT GGGCGAGACC
 151  GGCATGGATA TGGGGGCGGT ATTCGTCGCT ACCTGTATCG CGTCTGCCAT
 201  CGGCTGTTTT GTTATGGGTT TTGTCGGCAA CTATCCGATT GCACTCGCAC
 251  CGGGGATGGG GCTGAATGCC TATTTCACCT TTGCCGTCGT TAAGGGTATG
 301  GGCGTGCCTT GGCAGGTTGC GTTGGGTGCG GTGTTCATCT CCGGTCTGAT
 351  TTTCATCCTG TTCAGCTTTT TTAAAGTCAG GGAAATGCTG GTCAACGCAC
 401  TGCCTATGGG TTTGAAAATG TCGATTGCTG CCGGTATCGG TTTGTTTTTG
 451  GCACTGATTT CCCTGAAAGG CGCAGGCATT ATCGTTGCCA ATCCGGCAAC
 501  CTTGGTCGGC TTGGGCGATA TTCATCAGCC GTCCGCGTTG TTGGCACTGT
 551  TCGGTTTTGC CATGGTGGTC GTATTGGGAC ATTTCCGCGT TCAAGGCGCA
 601  ATCATCATCA CCATTTTGAC GATTACCGTC ATTGCCAGCC TGATGGGTTT
 651  GAACGAATTT CACGGCATCA TCGGCGAAGT GCCGAGCATT GCGCCGACTT
 701  TTATGCAGAT GGATTTTAAA GGGTTGTTTA CCGTCAGCAT GGTCAGCGTG
 751  ATTTTCGTCT TTTTCCTAGT CGATCTGTTC GACAGTACCG GAACACTGGT
 801  CGGTGTATCG CATCGTGCCG GACTGCTGGT GGACGGTAAG CTGCCCCGCC
 851  TGAAACGCGC ACTGCTTGCA GACTCTACCG CTATTGTGGC AGGTGCGGCT
 901  TTGGGTACTT CTTCAACCAC GCCTTATGTG GAAAGTGCGG CGGGCGTATC
 951  GGCAGGCGGG CGGACAGGTC TGACGGCGGT TACCGTCGGC GTATTGATGC
1001  TCGCCTGCCT GATGTTTTCA CCTTTGGCGA AAAGTGTTCC CGCTTTTGCC
1051  ACCGCGCCCG CCCTGCTTTA [illegible]TCGGCACG CAGATGCTCC GCAGTGCGAG
1101  GGACATCGAT TGGGACGATA TGACGGAAGC CGCACCCGCA TTCCTGACCA
1151  TTGTCTTCAT GCCGTTTACC TATTCGATTG CAGACGGCAT CGCTTTCGGC
1201  TTCATCAGTT ATGCCGTGGT TAAACTTTTA TGCCGCCGCA CCAAAGACGT
1251  TCCGCCTATG GTATGGATTG TTGCCGTATT GTGGGCACTG AAATTCTGGT
```

```
1301  ATTTGGGCTG A
```

This corresponds to the amino acid sequence <SEQ ID 344; ORF 097.a>:

a097.pep

```
  1  MDTSKQTLLD GIFKLKANGT TVRTELMAGL TTFLTMCYIV IVNPLILGET
 51  GMDMGAVFVA TCIASAIGCF VMGFVGNYPI ALAPGMGLNA YFTFAVVKGM
101  GVPWQVALGA VFISGLIFIL FSFFKVREML VNALPMGLKM SIAAGIGLFL
151  ALISLKGAGI IVANPATLVG LGDIHQPSAL LALFGFAMVV VLGHFRVQGA
201  IIITILTITV IASLMGLNEF HGIIGEVPSI APTFMQMDFK GLFTVSMVSV
251  IFVFFLVDLF DSTGTLVGVS HRAGLLVDGK LPRLKRALLA DSTAIVAGAA
301  LGTSSTTPYV ESAAGVSAGG RTGLTAVTVG VLMLACLMFS PLAKSVPAFA
351  TAPALLYVGT QMLRSARDID WDDMTEAAPA FLTIVFMPFT YSIADGIAFG
401  FISYAVVKLL CRRTKDVPPM VWIVAVLWAL KFWYLG*
```

m097/a097 99.3% identity in 436 aa overlap

```
                     10        20        30        40        50        60
m097.pep    MDTSKQTLLDGIFKLKANGTTVRTELMAGLTTFLTMCYIVIVNPXILGETGMDMGAVFVA
            |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
a097        MDTSKQTLLDGIFKLKANGTTVRTELMAGLTTFLTMCYIVIVNPLILGETGMDMGAVFVA
                     10        20        30        40        50        60

                     70        80        90       100       110       120
m097.pep    TCIASAIGCFVMGFVGNYPIALAPGMGLNAYFTFAVVKGMGVPWQVALGAVFISGLIFIL
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a097        TCIASAIGCFVMGFVGNYPIALAPGMGLNAYFTFAVVKGMGVPWQVALGAVFISGLIFIL
                     70        80        90       100       110       120

                    130       140       150       160       170       180
m097.pep    FSFFKVREMLVNALPMGLKMSIAAGIGLFLALISLKGAGIIVANPATLVGLGDIHQPSAL
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a097        FSFFKVREMLVNALPMGLKMSIAAGIGLFLALISLKGAGIIVANPATLVGLGDIHQPSAL
                    130       140       150       160       170       180

                    190       200       210       220       230       240
m097.pep    LALFGFAMVVVLGHFRVQGAIIITILTITVIASLMGLNEFHGIIGEVPSIAPTFMQMDFE
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
a097        LALFGFAMVVVLGHFRVQGAIIITILTITVIASLMGLNEFHGIIGEVPSIAPTFMQMDFK
                    190       200       210       220       230       240

                    250       260       270       280       290       300
m097.pep    GLFTVSMVSVIFVFFLVDLFDSTGTLVGISHRAGLLVDGKLPRLKRALLADSTAIVAGAA
            ||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
a097        GLFTVSMVSVIFVFFLVDLFDSTGTLVGVSHRAGLLVDGKLPRLKRALLADSTAIVAGAA
                    250       260       270       280       290       300

                    310       320       330       340       350       360
m097.pep    LGTSSTTPYVESAAGVSAGGRTGLTAVTVGVLMLACLMFSPLAKSVPAFATAPALLYVGT
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a097        LGTSSTTPYVESAAGVSAGGRTGLTAVTVGVLMLACLMFSPLAKSVPAFATAPALLYVGT
                    310       320       330       340       350       360

                    370       380       390       400       410       420
m097.pep    QMLRSARDIDWDDMTEAAPAFLTIVFMPFTYSIADGIAFGFISYAVVKLLCRRTKDVPPM
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a097        QMLRSARDIDWDDMTEAAPAFLTIVFMPFTYSIADGIAFGFISYAVVKLLCRRTKDVPPM
                    370       380       390       400       410       420

                    430
m097.pep    VWIVAVLWALKFWYLGX
            ||||||||||||||||
a097        VWIVAVLWALKFWYLGX
                    430
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 345>:

```
g098.seq
     1  ATGACCGCCG ACGGTCTCTT CGTCGCTTTC AACTTCAATA CGTTTGCCGT
    51  TGTGCGAATA TTGATACCAG TACAGCAGGA TGCTGCCCAG GCTGGCGATC
   101  AGTTTGTCGG CGATGTCGCG CGCTTCGCTG TCGGGATGGC TTTCGCGTTC
   151  GGGATGAACG CAGCCGAGCA TGGACACGCC GGTACGCATC ACGTCCATCG
   201  GATGGGTATG TGCAGGCAGG CTTTCCAAAA CTTTAATCAC ACGGATAGGC
   251  AGGCCGCGCA TGGATTTGAG CTTGGTTTTA TAAGCGGCCA GCTCGAATTT
   301  GTTGGGCAGA TGGCCGTGAA TCAGCAAGTG TGCGACTTCT TCAAACTCGC
   351  ATTTTTGTGC CAAATTAGAA TGTCGTAA
```

This corresponds to the amino acid sequence <SEQ ID 346; ORF 098.ng>:

```
g098.pep
     1  MTADGLFVAF NFNTFAVVRI LIPVQQDAAQ AGDQFVGDVA RFAVGMAFAF
    51  GMNAAEHGHA GTHHVHRMGM CRQAFQNFNH TDRQAAHGFE LGFISGQLEF
   101  VGQMAVNQQV CDFFKLAFLC QIRMS*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 347>:

```
m098.seq
     1  ATGACCGCCG ATGGTCTCTT CGTCGCTTTC AACCTCAATG CGTTTGCCGT
    51  TGTGCGAATA TTGATACCAG TACAAGAGGA TGCTGCCGAG GCTGGCGATC
   101  AGTTTGTCGG CGATGTCGCG CGCTTCACTT TCCGGATGGC TTTCACGTTC
   151  AGGATGAACG CAGCCCAGCA TGGATACGCC GGTACGCATT ACGTCCATCG
   201  GATGGGTATG TGCAGGCAGG CTTTCCAAAA CTTTAATCAC ACGGATAGGC
   251  AGGCCGCGCA TGGATTTGAG CTTGGTTTTA TAAGCGGCCA GCTCGAATTT
   301  GTTGGGCAGA TGGCCGTGAA TCAGCAGGTG GGCGACTTCT TCAAACTCGC
   351  ATTTTTGTGC CAAATCAGAA TGTCGTAA
```

This corresponds to the amino acid sequence <SEQ ID 348; ORF 098>:

```
m098.pep
     1  MTADGLFVAF NLNAFAVVRI LIPVQEDAAE AGDQFVGDVA RFTFRMAFTF
    51  RMNAAQHGYA GTHYVHRMGM CRQAFQNFNH TDRQAAHGFE LGFISGQLEF
   101  VGQMAVNQQV GDFFKLAFLC QIRMS*
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

ORF 098 shows 89.6% identity over a 125 aa overlap with a predicted ORF (ORF 098.ng) from *N. gonorrhoeae:*

```
m098/g098

                       10        20        30        40        50        60
m098.pep    MTADGLFVAFNLNAFAVVRILIPVQEDAAEAGDQFVGDVARFTFRMAFTFRMNAAQHGYA
            ||||||||||:|:|||||||||||||:|||:||||||||||||:   |||:| ||||:||:|
g098        MTADGLFVAFNFNTFAVVRILIPVQQDAAQAGDQFVGDVARFAVGMAFAFGMNAAEHGHA
                       10        20        30        40        50        60

                       70        80        90       100       110       120
m098.pep    GTHYVHRMGMCRQAFQNFNHTDRQAAHGFELGFISGQLEFVGQMAVNQQVGDFFKLAFLC
            |||:|||||||||||||||||||||||||||||||||||||||||||||| |||||||||
g098        GTHHVHRMGMCRQAFQNFNHTDRQAAHGFELGFISGQLEFVGQMAVNQQVCDFFKLAFLC
                       70        80        90       100       110       120

m098.pep    QIRMSX
            ||||||
g098        QIRMSX
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 349>:

```
a098.seq
     1  ATGACCGCCG ATGGTCTCTT CGTCGCTTTC AACCTCAATG CGTTTGCCGT
    51  TGTGCGAATA TTGATACCAG TACAAGAGGA TGCTGCCGAG GCTGGCGATC
```

```
101  AGTTTGTCGG CGATGTCGCG CGCTTCACTT TCCGGATGGC TTTCACGTTC
151  AGGATGAACG CAGCCCAGCA TGGATACGCC GGTACGCATT ACGTCCATCG
201  GATGGGTATG TGCAGGCAGG CTTTCCAAAA CTTTAATCAC ACGGATAGGC
251  AGGCCGCGCA TGGATTTGAG CTTGGTTT:A TAAGCGGCCA GCTCGAATTT
301  GTTGGGCAGA TGGCCGTGAA TCAGCAGGTG GGCGACTTCT TCAAACTCGC
351  ATTTTTGTGC CAAATCAGAA TGTCGTAA
```

This corresponds to the amino acid sequence <SEQ ID 350; ORF 098.a>:

a098.pep

```
  1  MTADGLFVAF NLNAFAVVRI LIPVQEDAAE AGDQFVGDVA RFTFRMAFTF
 51  RMNAAQHGYA GTHYVHRMGM CRQAFQNFNH TDRQAAHGFE LGFISGQLEF
101  VGQMAVNQQV GDFFKLAFLC QIRMS*
```

m098/a098   100.0% identity in 125 aa overlap

```
                    10        20        30        40        50        60
m098.pep    MTADGLFVAFNLNAFAVVRILIPVQEDAAEAGDQFVGDVARFTFRMAFTFRMNAAQHGYA
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a098        MTADGLFVAFNLNAFAVVRILIPVQEDAAEAGDQFVGDVARFTFRMAFTFRMNAAQHGYA
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m098.pep    GTHYVHRMGMCRQAFQNFNHTDRQAAHGFELGFISGQLEFVGQMAVNQQVGDFFKLAFLC
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a098        GTHYVHRMGMCRQAFQNFNHTDRQAAHGFELGFISGQLEFVGQMAVNQQVGDFFKLAFLC
                    70        80        90       100       110       120

m098.pep    QIRMSX
            ||||||
a098        QIRMSX
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 351>:

g099.seq

```
   1  ATGCTGGGAC GCGCGTCCAT GATGCGCCTG CCCGATATTG TCGGCGTGGA
  51  GCTGACGGGC AAACGGCAGG CGGGCATTAC TGCCACAGAC ATCGTGTTGG
 101  CACTGACCGA ATTCTTGCGT AAAGAGCGCG TGGTCGGGGC GTTTGTCGAA
 151  TTTTTCGGCG AGGGCGCGAG AAGCCTGTCT ATCGGCGACC GCGCGACCAT
 201  TTCCAACATG ACGCCGGAGT TCGGCGCGAC TGCCGCCATG TTCGCCATCG
 251  ACGCGCAAAC TATTGATTAT TTGAAACTGA CCGGACGTGA CGACGCGCAG
 301  GTGAAATTGG TGGAAACCTA CGCCAAAACC GCAGGCTTAT GGGCAGGTGG
 351  CTTGAAAACC GCCGTTTATC CGCGCGTTTT GAAATTTGAT TTGAGCAGCG
 401  TAACGCGCAA TATGGCAGGC CCGAGCAACC CGCACGCGCG TTTTGCCACC
 451  GCCGATTTGG CGGCGAAAGG GCTGGCGAAG CCTTACGAAG AGCCTTCAGA
 501  CGGCCAAATG CCTGACGGTG CAGTGATTAT TGCCGCGATT ACTTCGTGTA
 551  CCAATACTTC CAACCCGCGC AACGTTGTCG CCGCCGCACT GTTGGCACGC
 601  AATGCCAACC GCCTCGGCTT GAAACGCAAA CCTTGGGTGA AATCTTCGTT
 651  TGCCCCGGGT TCAAAAGTAG CCGGAATCTA TTTGAAAGAA GCAGGCTTGT
 701  TGCCCGAAAT GGAAAAACTC GGCTTCGGTA TCGTCGCCTT CGCATGTACC
 751  ACCTGTAACG GCATGAgcgG CGCGCTcgaC CCGAAAATCC AACAAGAAAT
 801  CATCGACCGC GAtttgtacg cCACCGCCGT ATTGTCAGGC AACCGCAACT
 851  TCGACGGCCG TATCCATCCG TATGCGAAAC AGGCTTTCCT CGCTTCGCCT
 901  CCTTTGGTCG TTGCCTACGC ATTGGCAGGT ACCATCCGTT TCGATATTGA
 951  AAACGACGTA CTCGGCGTTG CAGACGGCCG CGAAATCCGC CTGAAAGATA
1001  TCTGGCCGAC AGACGAAGAA ATCGATGCCA TCGTTGCCGA ATATGTGAAA
1051  CCGCAACAAT TCCGCGACAT TTATATCCCG ATGTCCGACA CCGGCACAGC
1101  GCAAAAAGCA CCAAGCCCGC TGTACGACTG GCGACCGATG TCCACCTACA
1151  TCCGCCGTCC GCCCTATTGG GAAGGCGCAC TGGCAGGGGA ACGTACATTA
1201  AGAGGTATGC GTCCGCCGGC GATTTTGCCC GACAACATCA CCACCGACCA
1251  CATCTCgcca tCCAATGCGA TTTTGGCCGG cagTGCcgca ggtgaATATT
```

```
1301  TGGCGAAAAT GGGTTTGCCT GAAGAagaCT TCAACTCTTA CGCAACCCAC
1351  CGCGGCGACC ACTTGACCGC CCAACGCGCA ACCTTCGCCA ATCCGAAACT
1401  GTTTAACGAA ATGGTGAGAA ACGAAGACGG CAGCGTACGC CAAGGTtcgt
1451  tggcacgcgT tgaacCAGAA GGCCAAACCA TGCGCATGTG GGAAGCCATC
1501  GAAACCTATA TGAACCGCAA ACAGCCGCTT ATCATCATTG CCGGTGCGGA
1551  CTATGGTCAA GGCTCAAGCC GCGACTGGGC GGCGAAGGGC GTGCGGCTGG
1601  CGGGTGTGGA AGCCATCGCC GCCGAAGGTT TCGAGCGCAT CCACCGCACC
1651  AACCTCATCG GCATGGGCGT CTTGCCGCTG CAATTCAAAC CCGGCACCAA
1701  CCGCCATACC CTGCAACTGG ACGGTACGGA AACCTACGAC GTTGTCGGCG
1751  AACGCACACC GCGCTGCGGC CTGACCCTCG TGATTCACCG TAAAAACGGA
1801  GAAACCGTCG AAGTTCCGGT TACCTGCCGC CCCGATACCG CAGAAGAAGC
1851  ATTGGTATAT GAAGCCGGCG GCGTATTGCA ACGGTTTGCA CAGGACTTTT
1901  TGGAAGGGAA CGCGGCTTAG
```

This corresponds to the amino acid sequence <SEQ ID 352; ORF 099.ng>:

```
g099.pep
  1  MLGRASMMRL PDIVGVELTG KRQAGITATD IVLALTEFLR KERVVGAFVE
 51  FFGEGARSLS IGDRATISNM TPEFGATAAM FAIDAQTIDY LKLTGRDDAQ
101  VKLVETYAKT AGLWAGGLKT AVYPRVLKFD LSSVTRNMAG PSNPHARFAT
151  ADLAAKGLAK PYEEPSDGQM PDGAVIIAAI TSCTNTSNPR NVVAAALLAR
201  NANRLGLKRK PWVKSSFAPG SKVAGIYLKE AGLLPEMEKL GFGIVAFACT
251  TCNGMSGALD PKIQQEIIDR DLYATAVLSG NRNFDGRIHP YAKQAFLASP
301  PLVVAYALAG SIRFDIENDV LGVADGREIR LKDIWPTDEE IDAIVAEYVK
351  PQQFRDIYIP MSDTGTAQKA PSPLYDWRPM STYIRRPPYW EGALAGERTL
401  RGMRPPAILP DNITTDHISP SNAILAGSAA GEYLAKMGLP EEDFNSYATH
451  RGDHLTAQRA TFANPKLFNE MVRNEDGSVR QGSLARVEPE GQTMRMWEAI
501  ETYMNRKQPL IIIAGADYGQ GSSRDWAAKG VRLAGVEAIA AEGFERIHRT
551  NLIGMGVLPL QFKPGTNRHT LQLDGTETYD VVGERTPRCG LTLVIHRKNG
601  ETVEVPVTCR PDTAEEALVY EAGGVLQRFA QDFLEGNAA*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 353>:

```
m099.seq
   1  ATGCTGGGAC GCGCGTCCAT GATGCGCCTG CCCGATATTG TCGGCGTTGA
  51  GCTGAACGGC AAACGGCAGG CGGGCATTAC GGCGACGGAT ATTGTGTTGG
 101  CACTGACCGA GTTTCTGCGC AAAGAACGCG TGGTCGGGGC GTTTGTCGAA
 151  TTCTTCGGCG AGGGCGCGAG AAGCCTGTCT ATCGGCGACC GCGCGACCAT
 201  TTCCAACATG ACGCCGGAGT TCGGCGCGAC TGCCGCGATG TTCGCTATTG
 251  ATGAGCAAAC CATTGATTAT TTGAAACTGA CCGGACGCGA CGACGCGCAG
 301  GTGAAATTGG TGGAAACCTA CGCCAAAACC GCAGGCTTGT GGGCAGATGC
 351  CTTGAAAACC GCCGTTTATC CTCGCGTTTT GAAATTTGAT TTGAGCAGCG
 401  TAACGCGCAA TATGGCAGGC CCAAGTAACC CGCATGCCCG TTTTGCGACC
 451  GCCGATTTGG CGGCGAAAGG GCTGGCGAAG CCTTACGAAG AGCCTTCGGA
 501  CGGCCAAATG CCCGACGGCT CGGTCATCAT CGCCGCGATT ACCAGTTGCA
 551  CCAACACTTC CAACCCGCGC AACGTTGTTG CCGCCGCGCT CTTGGCACGC
 601  AATGCCAACC GTCTCGGCTT GAAACGCAAA CCTTGGGTGA AATCTTCGTT
 651  TGCCCCGGGT TCAAAAGTAG CCGAAATCTA TTTGAAAGAA GCGGGCCTGT
 701  TGCCCGAAAT GGAAAAACTC GGCTTCGGTA TCGTCGCCTT CGCCTGCACC
 751  ACCTGCAACG GCATGAGTGG CGCGCTGGAT CCGAAAATCC AGAAAGAAAT
 801  CATCGACCGC GATTTGTACG CCACCGCCGT ATTATCAGGC AACCGCAACT
 851  TCGACGGCCG TATCCACCCG TATGCGAAAC AGGCTTTCCT CGCTTCGCCT
 901  CCGTTGGTCG TTGCCTACGC GCTGGCAGGC AGTATCCGTT TCGATATTGA
 951  AAACGACGTA CTCGGCGTTG CAGACGGCAA GGAAATCCGC CTGAAAGACA
1001  TTTGGCCTGC CGATGAAGAA ATCGATGCCG TCGTTGCCGA ATATGTGAAA
1051  CCGCAGCAGT TCCGCGATGT GTATGTACCG ATGTTCGACA CCGGCACAGC
1101  GCAAAAAGCA CCCAGTCCGC TGTACGATTG GCGTCCGATG TCCACCTACA
1151  TCCGCCGTCC GCCTTACTGG GAAGGCGCGC TGGCAGGGGA ACGCACATTA
1201  AGAGGTATGC GTCCGCTGGC GATTTTGCCC GACAACATCA CCACCGACCA
1251  CCTCTCGCCG TCCAATGCGA TTTTGGCCGT CAGTGCCGCA GGCGAGTATT
1301  TGGCGAAAAT GGGTTTGCCT GAAGAAGACT TCAACTCTTA CGCAACCCAC
1351  CGCGGCGACC ACTTGACCGC CCAACGCGCT ACCTTCGCCA ATCCGAAACT
1401  GTTTAACGAA ATGGTGAAAA ACGAAGACGG CAGCGTGCGC CAAGGCTCGT
1451  TCGCCCGCGT CGAACCCGAA GGCGAAACCA TGCGCATGTG GGAAGCCATC
1501  GAAACCTATA TGAACCGCAA ACAGCCGCTC ATCATCATTG CCGGTGCGGA
```

```
1551  CTATGGTCAA G^^^CAAGCC GCGACTGGGC TGCAAAAGGC GTACGCCTCG
1601  CCGGCGTAGA AGCGATTGTT GCCGAAGGCT TCGAGCGTAT CCACCGCACC
1651  AACCTTATCG GCATGGGCGT GTTGCCGCTG CAGTTCAAAC CCGACACCAA
1701  CCGCCATACC CTGCAACTGG ACGGTACGGA AACCTACGAC GTGGTCGGCG
1751  AACGCACACC GCGCTGCGAC CTGACCCTCG TGATTCACCG TAAAAACGGC
1801  GA^ACCGTTG AAGTTCCCGT TACCTGCTGC CTCGATACTG CAGAAGAAGT
1851  ATTGGTATAT GAAGCCGGCG GCGTGTTGCA ACGGTTTGCA CAGGATTTTT
1901  TGGAAGGGAA CGCGGCTTAG
```

This corresponds to the amino acid sequence <SEQ ID 354; ORF 099>:

```
m099.pep
   1  MLGRASMMRL PDIVGVELNG KRQAGITATD IVLALTEFLR KERVVGAFVE
  51  FFGEGARSLS IGDRATISNM TPEFGATAAM FAIDEQTIDY LKLTGRDDAQ
 101  VKLVETYAKT AGLWADALKT AVYPRVLKFD LSSVTRNMAG PSNPHARFAT
 151  ADLAAKGLAK PYEEPSDGQM PDGSVIIAAI TSCTNTSNPR NVVAAALLAR
 201  NANRLGLKRK PWVKSSFAPG SKVAEIYLKE AGLLPEMEKL GFGIVAFACT
 251  TCNGMSGALD PKIQKEIIDR DLYATAVLSG NRNFDGRIHP YAKQAFLASP
 301  PLVVAYALAG SIRFDIENDV LGVADGKEIR LKDIWPADEE IDAVVAEYVK
 351  PQQFRDVYVP MFDTGTAQKA PSPLYDWRPM STYIRRPPYW EGALAGERTL
 401  RGMRPLAILP DNITTDHLSP SNAILAVSAA GEYLAKMGLP EEDFNSYATH
 451  RGDHLTAQRA TFANPKLFNE MVKNEDGSVR QGSFARVEPE GETMRMWEAI
 501  ETYMNRKQPL IIIAGADYGQ GSSRDWAAKG VRLAGVEAIV AEGFERIHRT
 551  NLIGMGVLPL QFKPDTNRHT LQLDGTETYD VVGERTPRCD LTLVIHRKNG
 601  ETVEVPVTCC LDTAEEVLVY EAGGVLQRFA QDFLEGNAA*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 099 shows 96.2% identity over a 639 aa overlap with a predicted ORF (ORF 099.ng) from *N. gonorrhoeae:*

```
m099/g099

                    10        20        30        40        50        60
m099.pep    MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
            ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
g099        MLGRASMMRLPDIVGVELTGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m099.pep    IGDRATISNMTPEFGATAAMFAIDEQTIDYLKLTGRDDAQVKLVETYAKTAGLWADALKT
            ||||||||||||||||||||||| |||||||||||||||||||||||||||||||| :||
g099        IGDRATISNMTPEFGATAAMFAIDAQTIDYLKLTGRDDAQVKLVETYAKTAGLWAGGLKT
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m099.pep    AVYPRVLKFDLSSVTRNMAGPSNPHARFATADLAAKGLAKPYEEPSDGQMPDGSVIIAAI
            |||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
g099        AVYPRVLKFDLSSVTRNMAGPSNPHARFATADLAAKGLAKPYEEPSDGQMPDGAVIIAAI
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m099.pep    TSCTNTSNPRNVVAAALLARNANRLGLKRKPWVKSSFAPGSKVAEIYLKEAGLLPEMEKL
            |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
g099        TSCTNTSNPRNVVAAALLARNANRLGLKRKPWVKSSFAPGSKVAGIYLKEAGLLPEMEKL
                   190       200       210       220       230       240

                   250       260       270       280       290       300
m099.pep    GFGIVAFACTTCNGMSGALDPKIQKEIIDRDLYATAVLSGNRNFDGRIHPYAKQAFLASP
            ||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||
g099        GFGIVAFACTTCNGMSGALDPKIQQEIIDRDLYATAVLSGNRNFDGRIHPYAKQAFLASP
                   250       260       270       280       290       300

                   310       320       330       340       350       360
```

```
m099.pep    PLVVAYALAGSIRFDIENDVLGVADGKEIRLKDIWPADEEIDAVVAEYVKPQQFRDVYVP
            ||||||||||||||||||||||||||:|||||||||:||||||:||||||||||||:|:|
g099        PLVVAYALAGSIRFDIENDVLGVADGREIRLKDIWPTDEEIDAIVAEYVKPQQFRDIYIP
                310       320       330       340       350       360

                370       380       390       400       410       420
m099.pep    MFDTGTAQKAPSPLYDWRPMSTYIRRPPYWEGALAGERTLRGMRPLAILPDNITTDHLSP
            | |||||||||||||||||||||||||||||||||||||||||||| |||||||||||:||
g099        MSDTGTAQKAPSPLYDWRPMSTYIRRPPYWEGALAGERTLRGMRPPAILPDNITTDHISP
                370       380       390       400       410       420

                430       440       450       460       470       480
m099.pep    SNAILAVSAAGEYLAKMGLPEEDFNSYATHRGDHLTAQRATFANPKLFNEMVKNEDGSVR
            |||||| ||||||||||||||||||||||||||||||||||||||||||||||:||||||
g099        SNAILAGSAAGEYLAKMGLPEEDFNSYATHRGDHLTAQRATFANPKLFNEMVRNEDGSVR
                430       440       450       460       470       480

                490       500       510       520       530       540
m099.pep    QGSFARVEPEGETMRMWEAIETYMNRKQPLIIIAGADYGQGSSRDWAAKGVRLAGVEAIV
            |||:|||||||:|||||||||||||||||||||||||||||||||||||||||||||||:
g099        QGSLARVEPEGQTMRMWEAIETYMNRKQPLIIIAGADYGQGSSRDWAAKGVRLAGVEAIA
                490       500       510       520       530       540

                550       560       570       580       590       600
m099.pep    AEGFERIHRTNLIGMGVLPLQFKPDTNRHTLQLDGTETYDVVGERTPRCDLTLVIHRKNG
            ||||||||||||||||||||||||| ||||||||||||||||||||||| ||||||||||
g099        AEGFERIHRTNLIGMGVLPLQFKPGTNRHTLQLDGTETYDVVGERTPRCGLTLVIHRKNG
                550       560       570       580       590       600

                610       620       630       640
m099.pep    ETVEVPVTCCLDTAEEVLVYEAGGVLQRFAQDFLEGNAAX
            ||||||||  |||||:||||||||||||||||||||||||
g099        ETVEVPVTCRPDTAEEALVYEAGGVLQRFAQDFLEGNAAX
                610       620       630       640
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 355>:

a099.seq

```
   1  ATGCTGGGAC GCGCGTCCAT GATGCGCCTG CCCGATATTG TCGGCGTTGA
  51  GCTGAACGGC AAACGGAAGG CGGGCATTAC GGCGACGGAT ATTGTGTTGG
 101  CACTGACCGA GTTTCTGCGC AAAGAACGCG TGGTCGGGGC GTTTGTCGAA
 151  TTCTTCGGCG AGGGCGCGAG AAGCCTGTCT ATCGGCGACC GCGCGACCAT
 201  TTCCAACATG ACGCCGGAGT TCGGCGCGAC TGCCGCGATG TTCGCTATTG
 251  ATGAGCAAAC CATTGATTAT TTGAAACTGA CCGGACGCGA CGACGCGCAG
 301  GTGAAATTGG TGGAAACCTA CGCCAAAACC GCAGGCTTGT GGGCAGATGC
 351  CTTGAAAACC GCCGTTTATC CGCGCGTTTT GAAATTTGAT TTGAGCAGCG
 401  TAACGCGCAA TATGGCAGGC CCGAGCAACC CGCACGCGCG TTTTGCGACC
 451  GCCGATTTGG CCGGCAAAGG CTTGGCTAAA CCTTACGAAG AGCCTTCAGA
 501  CGGCCAAATG CCTGACGGTG CAGTGATTAT TGCCGCGATT ACTTCCTGTA
 551  CCAATACTTC CAATCCGCGC AACGTTGTCG CCGCCGCGCT GTTGGCACGC
 601  AATGCCAACC GCCTCGGCTT GCAACGCAAA CCTTGGGTGA AATCTTCGTT
 651  TGCCCCGGGT TCAAAAGTAG CCGAAATCTA TTTGAAAGAA GCAGATCTGC
 701  TGCCCGAAAT GGAAAAACTC GGCTTCGGTA TCGTTGCCTT CGCATGTACC
 751  ACCTGTAACG GCATGAGCGG CGCGCTGGAT CCGAAAATCC AGAAAGAAAT
 801  CATCGACCGC GATTTGTACG CCACCGCCGT ATTGTCAGGC AACCGCAACT
 851  TTGACGGCCG TATCCATCCG TATGCGAAAC AGGCTTTCCT CGCTTCGCCT
 901  CCGTTGGTCG TTGCCTACGC GCTGGCAGGC AGCATCCGTT TCGATATTGA
 951  AAACGACGTA CTCGGCGTTG CAGACGGCAA AGAAATCCGC CTGAAAGACA
1001  TTTGGCCTAC CGATGAAGAA ATCGATGCCA TCGTTGCCGA ATATGTGAAA
1051  CCGCAGCAAT TTCGCGACGT TTATATCCCG ATGTTCGACA CCGGCACAGC
1101  GCAAAAAGCA CCAAGCCCGC TGTACGACTG GCGTCCAATG TCTACCTATA
1151  TCCGCCGCCC ACCTTACTGG GAAGGCGCAC TGGCAGGGGA ACGCACATTA
1201  AGCGGTATGC GTCCGCTGGC GATTTTGCCC GACAACATCA CCACCGACCA
```

```
1251  TCTCTCGCCA TCCAATGCGA TTTTGGCAAG CAGTGCCGCA GGCGAATATT
1301  TGGCAAAAAT GGGTTTGCCT GAAGAAGACT TCAACTCTTA CGCAACCCAC
1351  CGTGGCGACC ACTTGACCGC CCAACGCGCA ACCTTCGCCA ATCCGAAACT
1401  GTTTAACGAA ATGGTGAGAA ACGAAGACGG CAGCGTACGC CAAGGTTCGC
1451  TGGCACGCGT TGAACCCGAA GGCCAAACCA TGCGCATGTG GGAAGCCATC
1501  GAAACCTATA TGAACCGCAA ACAGCCGCTC ATCATCATTG CCGGCGCGGA
1551  CTACGGTCAA GGCTCAAGCC GCGACTGGGC TGCAAAAGGC GTACGCCTCG
1601  CCGGCGTGGA AGCGATTGTT GCCGAAGGCT TCGAGCGTAT CCACCGCACC
1651  AACTTGATCG GTATGGGCGT GTTGCCGCTG CAGTTCAAAC CGGGTACCAA
1701  CCGCCACACC CTGCAACTGG ACGGTACGGA AACCTACGAC GTTGTCGGCG
1751  AACGCACACC GCGCTGCGAC CTGACCCTTG TGATTCACCG TAAAAACGGC
1801  GAGACCGTCG AAGTCCCCAT TACCTGCCGC CTCGATACCG CAGAAGAAGT
1851  GTTGGTATAT GAAGCCGGTG GCGTATTGCA ACGGTTTGCA CAGGATTTTT
1901  TGGAAGGGAA CGCGGCTTAG
```

This corresponds to the amino acid sequence <SEQ ID 356; ORF 099.a>:

a099.pep

```
  1  MLGRASMMRL PDIVGVELNG KRKAGITATD IVLALTEFLR KERVVGAFVE
 51  FFGEGARSLS IGDRATISNM TPEFGATAAM FAIDEQTIDY LKLTGRDDAQ
101  VKLVETYAKT AGLWADALKT AVYPRVLKFD LSSVTRNMAG PSNPHARFAT
151  ADLAGKGLAK PYEEPSDGQM PDGAVIIAAI TSCTNTSNPR NVVAAALLAR
201  NANRLGLQRK PWVKSSFAPG SKVAEIYLKE ADLLPEMEKL GFGIVAFACT
251  TCNGMSGALD PKIQKEIIDR DLYATAVLSG NRNFDGRIHP YAKQAFLASP
301  PLVVAYALAG SIRFDIENDV LGVADGKEIR LKDIWPTDEE IDAIVAEYVK
351  PQQFRDVYIP MFDTGTAQKA PSPLYDWRPM STYIRRPPYW EGALAGERTL
401  SGMRPLAILP DNITTDHLSP SNAILASSAA GEYLAKMGLP EEDFNSYATH
451  RGDHLTAQRA TFANPKLFNE MVRNEDGSVR QGSLARVEPE GQTMRMWEAI
501  ETYMNRKQPL IIIAGADYGQ GSSRDWAAKG VRLAGVEAIV AEGFERIHRT
551  NLIGMGVLPL QFKPGTNRHT LQLDGTETYD VVGERTPRCD LTLVIHRKNG
601  ETVEVPITCR LDTAEEVLVY EAGGVLQRFA QDFLEGNAA*
```

m099/a099   97.5% identity in 639 aa overlap

```
                    10        20        30        40        50        60
m099.pep    MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
            ||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
a099        MLGRASMMRLPDIVGVELNGKRKAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m099.pep    IGDRATISNMTPEFGATAAMFAIDEQTIDYLKLTGRDDAQVKLVETYAKTAGLWADALKT
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a099        IGDRATISNMTPEFGATAAMFAIDEQTIDYLKLTGRDDAQVKLVETYAKTAGLWADALKT
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m099.pep    AVYPRVLKFDLSSVTRNMAGPSNPHARFATADLAAKGLAKPYEEPSDGQMPDGSVIIAAI
            ||||||||||||||||||||||||||||||||||:||||||||||||||||||:||||||
a099        AVYPRVLKFDLSSVTRNMAGPSNPHARFATADLAGKGLAKPYEEPSDGQMPDGAVIIAAI
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m099.pep    TSCTNTSNPRNVVAAALLARNANRLGLKRKPWVKSSFAPGSKVAEIYLKEAGLLPEMEKL
            |||||||||||||||||||||||||||:||||||||||||||||||||||| ||||||||
a099        TSCTNTSNPRNVVAAALLARNANRLGLQRKPWVKSSFAPGSKVAEIYLKEADLLPEMEKL
                   190       200       210       220       230       240

                   250       260       270       280       290       300
m099.pep    GFGIVAFACTTCNGMSGALDPKIQKEIIDRDLYATAVLSGNRNFDGRIHPYAKQAFLASP
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a099        GFGIVAFACTTCNGMSGALDPKIQKEIIDRDLYATAVLSGNRNFDGRIHPYAKQAFLASP
                   250       260       270       280       290       300

                   310       320       330       340       350       360
m099.pep    PLVVAYALAGSIRFDIENDVLGVADGKEIRLKDIWPADEEIDAVVAEYVKPQQFRDVYVP
```

```
               ||||||||||||||||||||||||||||||||||||||:|||||:|||||||||||||:|
a099           PLVVAYALAGSIRFDIENDVLGVADGKEIRLKDIWPTDEEIDAIVAEYVKPQQFRDVYIP
                    310       320       330       340       350       360

                    370       380       390       400       410       420
m099.pep       MFDTGTAQKAPSPLYDWRPMSTYIRRPPYWEGALAGERTLRGMRPLAILPDNITTDHLSP
               |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
a099           MFDTGTAQKAPSPLYDWRPMSTYIRRPPYWEGALAGERTLSGMRPLAILPDNITTDHLSP
                    370       380       390       400       410       420

                    430       440       450       460       470       480
m099.pep       SNAILAVSAAGEYLAKMGLPEEDFNSYATHRGDHLTAQRATFANPKLFNEMVKNEDGSVR
               |||||| |.|||||||||||||||||||||:||||||||||||||||||||:|||||||
a099           SNAILASSAAGEYLAKMGLPEEDFNSYATHRGDHLTAQRATFANPKLFNEMVRNEDGSVR
                    430       440       450       460       470       480

                    490       500       510       520       530       540
m099.pep       QGSFARVEPEGETMRMWEAIETYMNRKQPLIIIAGADYGQGSSRDWAAKGVRLAGVEAIV
               |||:|||||||:||||||||||||||||||||||||||||||||||||||||||||||||
a099           QGSLARVEPEGQTMRMWEAIETYMNRKQPLIIIAGADYGQGSSRDWAAKGVRLAGVEAIV
                    490       500       510       520       530       540

                    550       560       570       580       590       600
m099.pep       AEGFERIHRTNLIGMGVLPLQFKPDTNRHTLQLDGTETYDVVGERTPRCDLTLVIHRKNG
               ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
a099           AEGFERIHRTNLIGMGVLPLQFKPGTNRHTLQLDGTETYDVVGERTPRCDLTLVIHRKNG
                    550       560       570       580       590       600

                    610       620       630       640
m099.pep       ETVEVPVTCCLDTAEEVLVYEAGGVLQRFAQDFLEGNAAX
               ||||||:|| ||||||||||||||||||||||||||||||
a099           ETVEVPITCRLDTAEEVLVYEAGGVLQRFAQDFLEGNAAX
                    610       620       630       640
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 357>:

g102.seq

```
   1  AtgtCCGCCA AAactccgtc gctcttcggc ggcgcgatga Ttatcgccgg
  51  gaaggttatc ggcgcAGgta tgttccccaa ccccaccgcc aacttggggg
 101  acgggttaat aggctcgctg attgtgctgc tgtacacctg gtttccattc
 151  tcctccggcg ccctcatgat tttggaagtc aacacccata acCCccgagg
 201  ggcaAGtttt gacaccATGg tcAAagacct gctcgGaCGc ggctggaaca
 251  tcatcaacgg catcgccgtc gctttggTCc tatacggctc gacctacgcg
 301  tacattttag tcggcggtga cctGACCGCC AAAGGCAtcg GCAgCGCAGT
 351  AGGCGGCAAA ATTTCgctca CCGTCGGACA actcgtcttc tTCGGCATCC
 401  TCGCCTTTTG CGTATGGGCA TCCGCACGCT TGGTCGACCG CTTTACCGGC
 451  GTCCTCATCG GCGGCATGGT ATTAACCTTT ATTTGGGCAA CCGGCGGCCT
 501  GGTTGCCGAT GCCAAACCGT CCGTCCTCTT CGACACCCAA GCCCCCGTCG
 551  GCACCGGCTA CTGGATTTAC GCCGCCACCG CCCTGCCCGT CTGCCTCGCT
 601  TCCTTCGGCT TCCACGGCAA CGTTTCCAGC CTGCTCAAAT ACTTTAAAGG
 651  CGACGcgcCc aaagtGgCGA aATCcatctg gGcaggtaca ttggTTGCCt
 701  tggtaattta cgtccTCTgg caaaccgcca tCcaaagcaa ccTGCcgcgc
 751  aacgagttcg cCCCcgtgat tgccgccgag aggcaactCT CCGTCCTgaa
 801  tgaaacccTG tccaaattcg cccaaaccgg cgatatggat aAaatattgt
 851  ccctatttcc ctacatggca atcgccacct ccttttttagg cgTAACctta
 901  ggcctgtttg acaacatcgc cgacatcttc aaatggaacg acagtatgtc
 951  cgggcggggc accaaaaccg tcgcgctgaa cttcctgccg CCCCtgattt
1001  cctggctgct cctccccacc ggcttcttta ccgccattgg tgcgtccggc
1051  ctggcggcaa ccgtctggga ccaagGcatc atccccgcca tgctgctcta
1101  cgtttccccc caaaaaattG gcGcaggcaa gacttataAa gtttaCGGCG
1151  gcttgtggct gatgttagtc ttccttttct gcatcgccaa catcgccgca
1201  CAGGTATTGA GccaAatgGa ACtcgtCccc GTATTTAAAG GATAA
```

This corresponds to the amino acid sequence <SEQ ID 358; ORF 102.ng>:

g102.pep

```
  1  MSAKTPSLFG GAMIIAGKVI GAGMFPNPTA NLGDGLIGSL IVLLYTWFPF
 51  SSGALMILEV NTHNPRGASF DTMVKDLLGR GWNIINGIAV ALVLYGSTYA
101  YILVGGDLTA KGIGSAVGGK ISLTVGQLVF FGILAFCVWA SARLVDRFTG
151  VLIGGMVLTF IWATGGLVAD AKPSVLFDTQ APVGTGYWIY AATALPVCLA
201  SFGFHGNVSS LLKYFKGDAP KVAKSIWAGT LVALVIYVLW QTAIQSNLPR
251  NEFAPVIAAE RQLSVLNETL SKFAQTGDMD KILSLFPYMA IATSFLGVTL
301  GLFDNIADIF KWNDSMSGRG TKTVALNFLP PLISWLLLPT GFFTAIGASG
351  LAATVWDQGI IPAMLLYVSP QKIGAGKTYK VYGGLWLMLV FLFGIANIAA
401  QVLSQMELVP VFKG*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 359>:

m102.seq

```
   1  ATGCCCAACA AAACCCCTTC ACTGTTCGGC GGCGCGATGA TTATCGCCGG
  51  CACGGTCATC GGCGCAGGCA TGCTCGCCAA CCCGACCGCC ACATCCGGCG
 101  TATGGTTTAC CGGCTCGCTG GCCGTGTTGC TGTACACCTG GTTTTCTATG
 151  CTTTCCAGCG GCCTGATGAT TTTGGAAGTC AACACCCATT ATCCGCACGG
 201  CGCAAGTTTC GACACGATGG TCAAAGACCT GCTCGGACGC GGCTGGAACA
 251  TCATCAACGG CATCGCCGTC GCCTTCGTTT TATACCTGCT TACTTACGCT
 301  TATATCTTCG TCGGCGGCGA CCTGACCGCC AAAGGCTTAG GCAGCGCGGC
 351  AGGCGGCGAC GTTTCACTCA CCGTCGGACA ACTCGTCTTC TTCGGCATCC
 401  TCGCCTTTTG CGTATGGGCA TCCGCACGCT TGGTCGACCG CTTCACCGGC
 451  GTCCTTATCG GCGGCATGGT ATTGACCTTT ATTTGGGCGG CCGGCGGGCT
 501  GATTGCCGAT GCCAAGCCGT CCGTCCTCTT CGATACCCAA GCCCCCGCCG
 551  GCACAAACTA CTGGATTTAC GCCGCCACCG CCCTGCCCGT CTGCCTCGCT
 601  TCCTTCGGCT TCCACGGCAA CGTCTCCAGC CTGCTCAAAT ACTTTAAAGG
 651  CGACGCGCCC AAAGTGGCTA AATCCATCTG GACGGGCACA CTGATTGCGC
 701  TGGTAATTTA CGTCCTCTGG CAAACCGCCA TCCAAGGCAA CCTGCCGCGC
 751  AACGAGTTCG CCCCCGTCAT CGCCGCCGAA GGGCAAGTCT CCGTCCTCAT
 801  CGAAACCCTG TCCAAATTCG CCCAAACCGG CAATATGGAC AAAATATTGT
 851  CCCTGTTTTC CTATATGGCG ATCGCCACCT CGTTTTTAGG CGTAACGCTC
 901  GGACTCTTCG ACTACATCGC CGACATCTTC AAATGGAACG ACAGCATCTC
 951  CGGCCGCACC AAAACCGCCG CGCTGACCTT CCTGCCGCCC CTGATTTCCT
1001  GCCTGCTCTT CCCCACCGGC TTCGTTACCG CCATCGGCTA CGTCGGCCTG
1051  GCGGCAACCG TCTGGACAGG CATCATCCCC GCCATGCTGC TCTACCGTTC
1101  GCGCAAAAAA TTCGGCGCAG GCAAAACCTA TAAAGTTTAC GGCGGCTTGT
1151  GGCTGATGGT TTGGGTCTTC CTTTTCGGCA TCGTCAACAT CGCCGCACAG
1201  GTATTGAGCC AAATGGAACT CGTCCCCGTA TTTAAAGGAT AA
```

This corresponds to the amino acid sequence <SEQ ID 360; ORF 102>:

m102.pep..

```
  1  MPNKTPSLFG GAMIIAGTVI GAGMLANPTA TSGVWFTGSL AVLLYTWFSM
 51  LSSGLMILEV NTHYPHGASF DTMVKDLLGR GWNIINGIAV AFVLYLLTYA
101  YIFVGGDLTA KGLGSAAGGD VSLTVGQLVF FGILAFCVWA SARLVDRFTG
151  VLIGGMVLTF IWAAGGLIAD AKPSVLFDTQ APAGTNYWIY AATALPVCLA
201  SFGFHGNVSS LLKYFKGDAP KVAKSIWTGT LIALVIYVLW QTAIQGNLPR
251  NEFAPVIAAE GQVSVLIETL SKFAQTGNMD KILSLFSYMA IATSFLGVTL
301  GLFDYIADIF KWNDSISGRT KTAALTFLPP LISCLLFPTG FVTAIGYVGL
351  AATVWTGIIP AMLLYRSRKK FGAGKTYKVY GGLWLMVWVF LFGIVNIAAQ
401  VLSQMELVPV FKG*
```

m102/g102 86.0% identity in 415 aa overlap

```
                     10        20        30        40        50        60
m102.pep     MPNKTPSLFGGAMIIAGTVIGAGMLANPTATSGVWFTGSLAVLLYTWFSMLSSGLMILEV
             |  |||||||||||||| ||||||: ||||: |   :||| ||||||| : |::||||||
g102         MSAKTPSLFGGAMIIAGKVIGAGMFPNPTANLGDGLIGSLIVLLYTWFPFSSGALMILEV
                     10        20        30        40        50        60

                     70        80        90       100       110       120
m102.pep     NTHYPHGASFDTMVKDLLGRGWNIINGIAVAFVLYLLTYAYIFVGGDLTAKGLGSAAGGD
             ||| |:|||||||||||||||||||||||||:|||  |||||:|||||||||:|||:|| 
g102         NTHNPRGASFDTMVKDLLGRGWNIINGIAVALVLYGSTYAYILVGGDLTAKGIGSAVGGK
                     70        80        90       100       110       120

                    130       140       150       160       170       180
m102.pep     VSLTVGQLVFFGILAFCVWASARLVDRFTGVLIGGMVLTFIWAAGGLIADAKPSVLFDTQ
             :||||||||||||||||||||||||||||||||||||||||||:|||:||||||||||||
g102         ISLTVGQLVFFGILAFCVWASARLVDRFTGVLIGGMVLTFIWATGGLVADAKPSVLFDTQ
                    130       140       150       160       170       180

                    190       200       210       220       230       240
m102.pep     APAGTNYWIYAATALPVCLASFGFHGNVSSLLKYFKGDAPKVAKSIWTGTLIALVIYVLW
```

```
               ||:||:||||||||||||||||||||||||||||||||||||||||||||:|||:||||||
g102           APVGTGYWIYAATALPVCLASFGFHGNVSSLLKYFKGDAPKVAKSIWAGTLVALVIYVLW
                       190       200       210       220       230       240

                       250       260       270       280       290       300
m102.pep       QTAIQGNLPRNEFAPVIAAEGQVSVLIETLSKFAQTGNMDKILSLFCYMAIATSFLGVTL
               |||||:|||||||||||||| .:||| |||||||||:||||||| ||||||||||||
g102           QTAIQSNLPRNEFAPVIAAERQLSVLNETLSKFAQTGDMDKILSLFPYMAIATSFLGVTL
                       250       260       270       280       290       300

                       310       320       330       340       350
m102.pep       GLFDYIADIFKWNDSISGR-TKTAALTFLPPLISCLLFPTGFVTAIGYVGLAATVWT-GI
               |||| |||||||||:||| |||:||:||||||| ||:    | ||||  ||||||| ||
g102           GLFDNIADIFKWNDSMSGRGTKTVALNFLPPLISWLLLFTGFFTAIGASGLAATVWDQGI
                       310       320       330       340       350       360

            360       370       380       390       400       410
m102.pep       IPAMLLYRSRKKFGAGKTYKVYGGLWLMVWVFLFGIVNIAAQVLSQMELVPVFKGX
               ||||||| | :|:|||||||||||||||: ||||||:||||||||||||||||||
g102           IPAMLLYVSPQKIGAGKTYKVYGGLWLML-VFLFGIANIAAQVLSQMELVPVFKGX
                       370       380       390       400       410
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 361>:

a102.seq

```
   1  ATGCCCACCA AAACCCCTTC ACTGTTCGGC GGCGCGATGA TTATCGCCGG
  51  CACGNTCATC GGCGCAGGTA TGCTCGCCAA CCCGACCGCC ACATCCGGCG
 101  TATGGTTTAC CGGCTCGCTG GCCGTGTTGC TGTACACCTG GTTTTCCATG
 151  CTCTCCAGCG GCCTGATGAT TTTGGAAGTC AACACCCACT ACCCCCACGG
 201  CGCGANCTTC GACACCATGG TTAAAGACCT GCTCGGACGG AGCTGGAACA
 251  TCATCAACGG CATCGCCGTC GCCTTCGTTT TATACCTGCT TACTTACGCT
 301  TATATCTTCG TCGGCGGCGA CCTGACCGCC AAAGGCTTAG GCAGCGCGGC
 351  AGGCGGCAAT GTTTCACTCA CCGTCGGACA ACTCGTCTTC TTCGGCATTC
 401  TCGCCTTTTG CGTATGGGCA TCCGCACGCT TGGTCGACCG ATTCACCAGC
 451  GTCCTCATCG GCGGCATGGT ATTAACCTTT ATTTGGGCAA CCGGCGGCCT
 501  GATTGCCGAT GCCAAACTGC CCGTCCTCTT CGACACCCAA GCCCCTACCG
 551  GCACCAACTA CTGGATTTAT GTCGCCACCG CCCTGCCCGT CTGCCTTGCG
 601  TCATTCGGTT TCCACGGCAA CGTCTCCAGC CTGCTCAAAT ACTTTAAAGG
 651  CGACGCGCCC AAAGTGGCTA AATCCATCTG GACGGGCACA CTGATTGCGC
 701  TGGTAATTTA CGTCCTCTGG CAAACCGCCA TCCAANGCAA CCTGCCGCGC
 751  AACGAGTTCG CCCCCGTGAT TGCCGCCGAA GGGCAAGTCT CCGTCNTGAT
 801  TGAAACCCTG TCCAAATTCG CCCAAACCGG CAATATGGAC AAAATATTGT
 851  CCCTGTTTTC CTATATGGCG ATCGCCACCT CGTTTTTAGG CGTAACGCTC
 901  GGACTCTTCG ACTACATCGC CGACATCTTC AAATGGAACG ACAGCGTGTC
 951  CGGCCGCACC AAAACCGCCG CGCTGACCTT CCTGCCGCCT NTAATTTCCT
1001  GCCTGCTCTT CCCCACCGGC TTTGTTACCG CCATCGGNTA CGTCGGCCTG
1051  GCGGCAACCG TCTGGACAGG CATCATCCCC GCCATGCTGC TNTACCGTTC
1101  GCGCAAAAAA TTCGGCGCAG GCAAAACCTA TAAAGTTTAC GGCGGCTTGT
1151  GGCTGATGGT TTGGGTCTTC CTTTTCGGCA TCNTCAACAT CGCCGCACAN
1201  GTATTGAGCC AAATGGAACT CGTCCCCGTA TTTAAAGGAT AA
1202
```

This corresponds to the amino acid sequence <SEQ ID 362; ORF 102.a>:

a102.pep

```
   1  MPTKTPSLFG GAMIIAGTXI GAGMLANPTA TSGVWFTGSL AVLLYTWFSM
  51  LSSGLMILEV NTHYPHGAXF DTMVKDLLGR SWNIINGIAV AFVLYLLTYA
 101  YIFVGGDLTA KGLGSAAGGN VSLTVGQLVF FGILAFCVWA SARLVDRFTS
 151  VLIGGMVLTF IWATGGLIAD AKLPVLFDTQ APTGTNYWIY VATALPVCLA
 201  SFGFHGNVSS LLKYFKGDAP KVAKSIWTGT LIALVIYVLW QTAIQXNLPR
 251  NEFAPVIAAE GQVSVXIETL SKFAQTGNMD KILSLFSYMA IATSFLGVTL
 301  GLFDYIADIF KWNDSVSGRT KTAALTFLPP XISCLLFPTG FVTAIGYVGL
 351  AATVWTGIIP AMLLYRSRKK FGAGKTYKVY GGLWLMVWVF LFGIXNIAAX
 401  VLSQMELVPV FKG*
```

m102 / a102 95.9% identity in 413 aa overlap

```
                     10        20        30        40        50        60
m102.pep     MPNKTPSLFGGAMIIAGTVIGAGMLANPTATSGVWFTGSLAVLLYTWFSMLSSGLMILEV
             ||:||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
a102         MPTKTPSLFGGAMIIAGTXIGAGMLANPTATSGVWFTGSLAVLLYTWFSMLSSGLMILEV
                     10        20        30        40        50        60

                     70        80        90       100       110       120
m102.pep     NTHYPHGASFDTMVKDLLGRGWNIINGIAVAFVLYLLTYAYIFVGGDLTAKGLGSAAGGD
             |||||||| |||||||||||:||||||||||||||||||||||||||||||||||||||:
a102         NTHYPHGAXFDTMVKDLLGRSWNIINGIAVAFVLYLLTYAYIFVGGDLTAKGLGSAAGGN
                     70        80        90       100       110       120

                    130       140       150       160       170       180
m102.pep     VSLTVGQLVFFGILAFCVWASARLVDRFTGVLIGGMVLTFIWAAGGLIADAKPSVLFDTQ
             |||||||||||||||||||||||||||||:|||||||||||||:||||||||  ||||||
a102         VSLTVGQLVFFGILAFCVWASARLVDRFTSVLIGGMVLTFIWATGGLIADAKLPVLFDTQ
                    130       140       150       160       170       180

                    190       200       210       220       230       240
m102.pep     APAGTNYWIYAATALPVCLASFGFHGNVSSLLKYFKGDAPKVAKSIWTGTLIALVIYVLW
             ||:|||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
a102         APTGTNYWIYVATALPVCLASFGFHGNVSSLLKYFKGDAPKVAKSIWTGTLIALVIYVLW
                    190       200       210       220       230       240

                    250       260       270       280       290       300
m102.pep     QTAIQGNLPRNEFAPVIAAEGQVSVLIETLSKFAQTGNMDKILSLFSYMAIATSFLGVTL
             ||||| ||||||||||||||||||| ||||||||||||||||||||||||||||||||||
a102         QTAIQXNLPRNEFAPVIAAEGQVSVXIETLSKFAQTGNMDKILSLFSYMAIATSFLGVTL
                    250       260       270       280       290       300

                    310       320       330       340       350       360
m102.pep     GLFDYIADIFKWNDSISGRTKTAALTFLPPLISCLLFPTGFVTAIGYVGLAATVWTGIIP
             |||||||||||||||:|||||||||||||| |||||||||||||||||||||||||||||
a102         GLFDYIADIFKWNDSVSGRTKTAALTFLPPXISCLLFPTGFVTAIGYVGLAATVWTGIIP
                    310       320       330       340       350       360

                    370       380       390       400       410
m102.pep     AMLLYRSRKKFGAGKTYKVYGGLWLMVWVFLFGIVNIAAQVLSQMELVPVFKGX
             |||||||||||||||||||||||||||||||||| |||| ||||||||||||||
a102         AMLLYRSRKKFGAGKTYKVYGGLWLMVWVFLFGIXNIAAXVLSQMELVPVFKGX
       370   380   390   400   410
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 363>:

```
g105.seq
      1  Atgtccgcag aaaCATACAc acAAAtcggc tGGgtaggct taggGcaaat
     51  gGgtctgcct atgGTAACGC GGCTCTTGGA CGGCGGCATC GAAGTCGGCG
    101  TATACAACCG CTCGCCCGAC AAAACTGCCC CCATCTCcgc CAAAGGAGCA
    151  AAAGTTTACG GCagcACCGC CGAACTCGTC CGCGCCTGCC CCGTCATTTT
    201  CCTGATGGTT TCCGACTATG CCGCCGTGTG CGACATCCTG AACGGAGTCC
    251  GCGACGGATT GGCCGGCAAA ATCATCGTCA ACATGAGCAC CATCTCCCCG
    301  ACCGAAAACC TCGCCGTCAA AGCACTTGTC GAAGCCGCAG GCGGACAGTT
    351  TGCCGAAGCA CCCGTTTCCG GATCGGTCGG ACCCGCCACC AACGGCACAC
    401  TGCTGATTCT GTTCGGCGGC AGCGAAGCCG TTTTAAACCC GCTGCAAAAA
    451  ATATTTTCCC TTGTCGGCAA AAAAACCTTC CATTTCGGCG ATGTCGGCAA
    501  AGGCTCGGGC GCGAAACTCG TCTTGAACTC GCTCTTAGGC ATTTTCGGCG
    551  AAGCGTACAG CGAAGCGATG CTGATGGCGC GGCAGTTCGG CATCGATACC
    601  GACACCATCG TCGAAGCCAT CGGCGGCTCG GCAATGGACT CGCCTATGTT
    651  TCAAACAAAA AAATCACTAT GGGCAAACCG TGAGTTCCCC CCTGCCTTTG
    701  CACTCAAACA CGCTTCCAAA GAcctTAACC TCGccgtcAA AGAGCTTGAA
    751  CAGGCAGGCA ACACCCTGCC CGCCGTCGAA ACCGTTGCTG CCAGCTACCG
    801  CAAAGCAGTT GAAGCCGGCT ACGGCGAACA GGACGTTTCC GGCGTTTACC
    851  TGAAATTGGC AGAACACTGA
```

This corresponds to the amino acid sequence <SEQ ID 364; ORF 105.ng>:

```
g105.pep
     1  MSAETYTQIG WVGLGQMGLP MVTRLLDGGI EVGVYNRSPD KTAPISAKGA
    51  KVYGSTAELV RACPVIFLMV SDYAAVCDIL NGVRDGLAGK IIVNMSTISP
   101  TENLAVKALV EAAGGQFAEA PVSGSVGPAT NGTLLILFGG SEAVLNPLQK
   151  IFSLVGKKTF HFGDVGKGSG AKLVLNSLLG IFGEAYSEAM LMARQFGIDT
   201  DTIVEAIGGS AMDSPMFQTK KSLWANREFP PAFALKHASK DLNLAVKELE
   251  QAGNTLPAVE TVAASYRKAV EAGYGEQDVS GVYLKLAEH
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 365>:

```
m105.seq
     1  ATGTCCGCAA ACGAATACGC ACAAATCGGC TGGaTAGGCT TAGGGCAAAT
    51  GGGTCTGCCT ATGGTAACGC GGCTCTTGGA CGGCGGCATC GAAGTCGGCG
   101  TATACAACCG CTCGCCCGAC AAAACTGCCC CCATCTCCGC CAAAGGCGCA
   151  AAAGTTTACG GCAACACCGC CGAACTCGTC CGCGACTATC CCGTCATTTT
   201  CCTGATGGTT TCCGACTATG CCGCCGTGTG CGACATCCTG AACGGAGTCC
   251  GCGACGGATT GGCCGGCAAm ATCATCGTCA ACATGAGCAC CATCTCCCCG
   301  ACCGAAAaGC TCGCCGTCAA AGCACTTGTC GAAGCGCAGm GaCAGTTTGC
   351  CGAAGCACCC GTTTCCGGAT CGGTCGGGCC CGCCACCAAC GGCACGCTGC
   401  TGATTCTGTT CGGCGGCAGC GAAcCGtTTT AAACCCGCTG CAAAAAATAT
   451  TTTCCCTCGT CCGCAAAAAA ACCTTCCATT TCGGCGATGT CGGCAAAGGT
   501  TCGGGCGCGA AACTCGTCTT GAACTCGCTC TTGGGCATTT TCGGCGAaCG
   551  TAcAGCGAAs GmTgCTGATG GCGCGGCAGT TCGGCATCGA TACCGACACC
   601  ATCGTCGAAG CCATCGGsGA CTCGGCAATG GACTCGCCCA TGTTCCAAAC
   651  CAAAAAATCC CTGTGGGCAA ACCGCGAATT CCCGmCCGmC TTCGCCCTCA
   701  AACACGCCTC CAAAGACCTC AACCTCGCCG TCAAAGAGCT TGAACAGGCA
   751  GGCAACACCC TGCCCGCCGT CGAAACCGTT GCTGCCAGCT ACCGCAAAGC
   801  AGTCGAAGCC GGCTACGGGA CACAGGACGT TTCCGGCGTT TACCTGAAAC
   851  TGGCAGAACA CTGA
```

This corresponds to the amino acid sequence <SEQ ID 366; ORF 105>:

```
m105.pep
     1  MSANEYAQIG WIGLGQMGLP MVTRLLDGGI EVGVYNRSPD KTAPISAKGA
    51  KVYGNTAELV RDYPVIFLMV SDYAAVCDIL NGVRDGLAGX IIVNMSTISP
   101  TEKLAVKALV EAQRQFAEAP VSGSVGPATN GTLLILFGGS EPFXTRCKKY
   151  FPSSAKKPSI SAMSAKVRAR NSSXTRSWAF SANVQRXXLM ARQFGIDTDT
   201  IVEAIGDSAM DSPMFQTKKS LWANREFPXX FALKHASKDL NLAVKELEQA
   251  GNTLPAVETV AASYRKAVEA GYGTQDVSGV YLKLAEH
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

ORF 105 shows 79.9% identity over a 289 aa overlap with a predicted ORF (ORF 105.ng) from *N. gonorrhoeae:*

```
m105/g105

                      10        20        30        40        50        60
g105.pep    MSAETYTQIGWVGLGQMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGSTAELV
            |||: |:||||:||||||||||||||||||||||||||||||||||||||||||:||||||
m105        MSANEYAQIGWIGLGQMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGNTAELV
                      10        20        30        40        50        60

                      70        80        90       100       110       120
g105.pep    RACPVIFLMVSDYAAVCDILNGVRDGLAGKIIVNMSTISPTENLAVKALVEAAGGQFAEA
            |  |||||||||||||||||||||||||| ||||||||||||:|||||||||   |||||
m105        RDYPVIFLMVSDYAAVCDILNGVRDGLAGXIIVNMSTISPTEKLAVKALVEAQR-QFAEA
                      70        80        90       100       110

                     130       140       150       160       170       180
g105.pep    PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFHFGDVGKGSGAKLVLNSLLG
            ||||||||||||||||||||||   :  :| |    :||    :: ::     |:   :
m105        PVSGSVGPATNGTLLILFGGSEPFXTRCKKYFPSSAKKP-SISAMSAKVRARNSSXTRSW
          120       130       140       150       160       170
```

```
                 190       200       210       220       230       240
g105.pep  IFGEAYSEAMLMARQFGIDTDTIVEAIGGSAMDSPMFQTKKSLWANREFPPAFALKHASK
          |:   ::  ||||||||||||||||| |||||||||||||||||||||  ||||||||
m105      AFSANVQRXXLMARQFGIDTDTIVEAIGDSAMDSPMFQTKKSLWANREFPXXFALKHASK
          180      190       200       210       220       230

                 250       260       270       280      289
g105.pep  DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAEH
          ||||||||||||||||||||||||||||||||||| ||||||||||||
m105      DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGTQDVSGVYLKLAEH
          240       250       260       270       280
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 367>:

a105.seq

```
  1  ATGTCCGCAA ACGAATACAC ACAAATCGGC TGGATAGGCT TAGGGCAAAT
 51  GGGTCTGCCT ATGGTAACGC GGCTCTTGGA CGGCGGCATC GAAGTCGGCG
101  TATACAACCG CTCGCCCGAC AAAACTGCCC CCATCTCCGC CAAAGGCGCA
151  AAAGTTTACG GCAACACCGC CGAACTCGTC CGCGACTATC CCGTCATTTT
201  CCTGATGGTT TCCGACTATG CCGCCGTGTG CGACATCCTG AACGGAGTCC
251  GCGACGGATT GGCCGGCAAA ATCATCGTCA ACATGAGCAC CATCTCCCCG
301  ACCGAAAACC TCGCCGTCAA AGCACTTGTC GAAGCCGCAG GCGGACAGTT
351  TGCCGAAGCA CCCGTTTCCG GATCGGTCGG GCCCGCCACC AACGGCACGC
401  TGCTGATTCT GTTCGGCGGC AGCGAAGCCG TTTTAAACCC GCTGCAAAAA
451  ATATTTTCCC TCGTCGGCAA AAAAACCTTC CATTTCGGCG ATGTCGGCAA
501  AGGTTCGGGC GCGAAACTCG TCTTGAACTC GCTCTTGGGC ATTTTCGGCG
551  AAGCGTACAG CGAAGCGATG CTGATGGCGC GGCAGTTCGG CATCGATACC
601  GACACCATCG TCGAAGCCAT CGGCGGCTCG GCAATGGACT CGCCCATGTT
651  CCAAACCAAA AAATCCCTGT GGGCAAACCG CGAATTCCCA CCCGCCTTCG
701  CCCTCAAACA CGCCTCCAAA GACCTCAACC TCGCCGTCAA AGAGCTTGAA
751  CAGGCAGGCA ACACCCTGCC CGCCGTCGAA ACCGTTGCTG CCAGCTACCG
801  CAAAGCAGTC GAAGCCGGCT ACGGCGAACA GGACGTTTCC GGCGTTTACC
851  TGAAATTGGC AGAACACTGA
```

This corresponds to the amino acid sequence <SEQ ID 368; ORF 105.a>:

a105.pep

```
  1  MSANEYTQIG WIGLGQMGLP MVTRLLDGGI EVGVYNRSPD KTAPISAKGA
 51  KVYGNTAELV RDYPVIFLMV SDYAAVCDIL NGVRDGLAGK IIVNMSTISP
101  TENLAVKALV EAAGGQFAEA PVSGSVGPAT NGTLLILFGG SEAVLNPLQK
151  IFSLVGKKTF HFGDVGKGSG AKLVLNSLLG IFGEAYSEAM LMARQFGIDT
201  DTIVEAIGGS AMDSPMFQTK KSLWANREFP PAFALKHASK DLNLAVKELE
251  QAGNTLPAVE TVAASYRKAV EAGYGEQDVS GVYLKLAEH*
```

m105/a105 96.5% identity in 289 aa overlap

```
                  10        20        30        40        50        60
m105.pep  MSANEYAQIGWIGLGQMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGNTAELV
          ||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
a105      MSANEYTQIGWIGLGQMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGNTAELV
                  10        20        30        40        50        60

                  70        80        90       100       110      119
m105.pep  RDYPVIFLMVSDYAAVCDILNGVRDGLAGKIIVNMSTISPTENLAVKALVEAAG-QFAEA
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
a105      RDYPVIFLMVSDYAAVCDILNGVRDGLAGKIIVNMSTISPTENLAVKALVEAAGGQFAEA
                  70        80        90       100       110       120

         120       130       140       150       160       170      179
m105.pep  PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFHFGDVGKGSGAKLVLNSLLG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a105      PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFHFGDVGKGSGAKLVLNSLLG
                 130       140       150       160       170       180

         180       190       200       210       220       230
```

```
m105.pep      IFGDV-QRXMLMARQFGIDTDTIVEAIGDSAMDSPMFQTKKSLWANREFPXAFALKHASK
              |||:: :: |||||||||||||||||||| |||||||||||||||||||||| |||||||||
a105          IFGEAYSEAMLMARQFGIDTDTIVEAIGGSAMDSPMFQTKKSLWANREFPPAFALKHASK
                      190       200       210       220       230       240

             240       250       260       270       280
m105.pep      DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAEHX
              ||||||||||||||||||||||||||||||||||||||||||||||||||
a105          DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAEHX
                      250       260       270       280       290
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 369>:

g105-1.seq

```
  1  ATGTCCGCAG AAACATACAC ACAAATCGGC TGGGTAGGCT TAGGGCAAAT
 51  GGGTCTGCCT ATGGTAACGC GGCTCTTGGA CGGCGGCATC GAAGTCGGCG
101  TATACAACCG CTCGCCCGAC AAAACTGCCC CCATCTCCGC CAAAGGAGCA
151  AAAGTTTACG GCAGCACCGC CGAACTCGTC CGCGCCTGCC CCGTCATTTT
201  CCTGATGGTT TCCGACTATG CCGCCGTGTG CGACATCCTG AACGGAGTCC
251  GCGACGGATT GGCCGGCAAA ATCATCGTCA ACATGAGCAC CATCTCCCCG
301  ACCGAAAACC TCGCCGTCAA AGCACTTGTC GAAGCCGCAG GCGGACAGTT
351  TGCCGAAGCA CCCGTTTCCG GATCGGTCGG ACCCGCCACC AACGGCACAC
401  TGCTGATTCT GTTCGGCGGC AGCGAAGCCG TTTTAAACCC GCTGCAAAAA
451  ATATTTTCCC TTGTCGGCAA AAAAACCTTC CATTTCGGCG ATGTCGGCAA
501  AGGCTCGGGC GCGAAACTCG TCTTGAACTC GCTCTTAGGC ATTTTCGGCG
551  AAGCGTACAG CGAAGCGATG CTGATGGCGC GGCAGTTCGG CATCGATACC
601  GACACCATCG TCGAAGCCAT CGGCGGCTCG GCAATGGACT CGCCCATGTT
651  TCAAACAAAA AAATCACTAT GGGCAAACCG TGAGTTCCCC CCTGCCTTTG
701  CACTCAAACA CGCTTCCAAA GACCTTAACC TCGCCGTCAA AGAGCTTGAA
751  CAGGCAGGCA ACACCCTGCC CGCCGTCGAA ACCGTTGCTG CCAGCTACCG
801  CAAAGCAGTT GAAGCCGGCT ACGGCGAACA GGACGTTTCC GGCGTTTACC
851  TGAAATTGGC AGAACACTGA
```

This corresponds to the amino acid sequence <SEQ ID 370; ORF 105-1.ng>:

g105-1.pep

```
  1  MSAETYTQIG WVGLGQMGLP MVTRLLDGGI EVGVYNRSPD KTAPISAKGA
 51  KVYGSTAELV RACPVIFLMV SDYAAVCDIL NGVRDGLAGK IIVNMSTISP
101  TENLAVKALV EAAGGQFAEA PVSGSVGPAT NGTLLILFGG SEAVLNPLQK
151  IFSLVGKKTF HFGDVGKGSG AKLVLNSLLG IFGEAYSEAM LMARQFGIDT
201  DTIVEAIGGS AMDSPMFQTK KSLWANREFP PAFALKHASK DLNLAVKELE
251  QAGNTLPAVE TVAASYRKAV EAGYGEQDVS GVYLKLAEH*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 371>:

m105-1.seq

```
  1  ATGTCCGCAA ACGAATACGC ACAAATCGGC TGGATAGGCT TAGGGCAAAT
 51  GGGTCTGCCT ATGGTAACGC GGCTCTTGGA CGGCGGCATC GAAGTCGGCG
101  TATACAACCG CTCGCCCGAC AAAACTGCCC CCATCTCCGC CAAAGGCGCA
151  AAAGTTTACG GCAACACCGC CGAACTCGTC CGCGACTATC CCGTCATTTT
201  CCTGATGGTT TCCGACTATG CCGCCGTGTG CGACATCCTG AACGGAGTCC
251  GCGACGGATT GGCCGGCAAA ATCATCGTCA ACATGAGCAC CATCTCCCCG
301  ACCGAAAACC TCGCCGTCAA AGCACTTGTC GAAGCCGCAG GCGGACAGTT
351  TGCCGAAGCA CCCGTTTCCG GATCGGTCGG GCCCGCCACC AACGGCACGC
401  TGCTGATTCT GTTCGGCGGC AGCGAAGCcG TTTTAAACCC GCTGCAAAAA
451  ATATTTTCCC TCGTCGGCAA AAAAACCTTC CATTTCGGCG ATGTCGGCAA
501  AGGTTCGGGC GCGAAACTCG TCTTGAACTC GCTCTTGGGC ATTTTCGGCG
551  AAGCGTACAG CGAAnCGATG CTGATGGCGC GGCAGTTCGG CATCGATACC
601  GACACCATCG TCGAAGCCAT CGGsGACTCG GCAATGGACT CGCCCATGTT
651  CCAAACCAAA AAATCCCTGT GGGCAAACCG CGAATTCCCG CCCGCCTTCG
701  CCCTCAAACA CGCCTCCAAA GACCTCAACC TCGCCGTCAA AGAGCTTGAA
751  CAGGCAGGCA ACACCCTGCC CGCCGTCGAA ACCGTTGCTG CCAGCTACCG
801  CAAAGCAGTC GAAGCCGGCT ACGGCGAACA GGACGTTTCC GGCGTTTACC
851  TGAAACTGGC AGAACACTGA
```

This corresponds to the amino acid sequence <SEQ ID 372; ORF 105-1>:

m105-1.pep

```
  1  MSANEYAQIG WIGLGQMGLP MVTRLLDGGI EVGVYNRSPD KTAPISAKGA
 51  KVYGNTAELV RDYPVIFLMV SDYAAVCDIL NGVRDGLAGK IIVNMSTISP
```

```
101  TENLAVKALV EAAGGQFAEA PVSGSVGPAT NGTLLILFGG SEAVLNPLQK
151  IFSLVGKKTF HFGDVGKGSG AKLVLNSLLG IFGEAYSEXM LMARQFGIDT
201  DTIVEAIGDS AMDSPMFQTK KSLWANREFP PAFALKHASK DLNLAVKELE
251  QAGNTLPAVE TVAASYRKAV EAGYGEQDVS GVYLKLAEH*
```

m105-1/g105-1 96.9% identity in 289 aa overlap

```
                    10        20        30        40        50        60
m105-1.pep  MSANEYAQIGWIGLGQMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGNTAELV
            |||: |:||||:||||||||||||||||||||||||||||||||||||||||||:|||||
g105-1      MSAETYTQIGWVGLGQMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGSTAELV
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m105-1.pep  RDYPVIFLMVSDYAAVCDILNGVRDGLAGKIIVNMSTISPTENLAVKALVEAAGGQFAEA
            |  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g105-1      RACPVIFLMVSDYAAVCDILNGVRDGLAGKIIVNMSTISPTENLAVKALVEAAGGQFAEA
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m105-1.pep  PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFHFGDVGKGSGAKLVLNSLLG
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g105-1      PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFHFGDVGKGSGAKLVLNSLLG
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m105-1.pep  IFGEAYSEXMLMARQFGIDTDTIVEAIGDSAMDSPMFQTKKSLWANREFPPAFALKHASK
            |||||||| |||||||||||||||||||| ||||||||||||||||||||||||||||||
g105-1      IFGEAYSEAMLMARQFGIDTDTIVEAIGGSAMDSPMFQTKKSLWANREFPPAFALKHASK
                   190       200       210       220       230       240

                   250       260       270       280       290
m105-1.pep  DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAEHX
            ||||||||||||||||||||||||||||||||||||||||||||||||||
g105-1      DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAEHX
                   250       260       270       280       290
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 373>:

a105-1.seq

```
  1  ATGTCCGCAA ACGAATACAC ACAAATCGGC TGGATAGGCT TAGGGCAAAT
 51  GGGTCTGCCT ATGGTAACGC GGCTCTTGGA CGGCGGCATC GAAGTCGGCG
101  TATACAACCG CTCGCCCGAC AAAACTGCCC CCATCTCCGC CAAAGGCGCA
151  AAAGTTTACG GCAACACCGC CGAACTCGTC CGCGACTATC CCGTCATTTT
201  CCTGATGGTT TCCGACTATG CCGCCGTGTG CGACATCCTG AACGGAGTCC
251  GCGACGGATT GGCCGGCAAA ATCATCGTCA ACATGAGCAC CATCTCCCCG
301  ACCGAAAACC TCGCCGTCAA AGCACTTGTC GAAGCCGCAG GCGGACAGTT
351  TGCCGAAGCA CCCGTTTCCG GATCGGTCGG GCCCGCCACC AACGGCACGC
401  TGCTGATTCT GTTCGGCGGC AGCGAAGCCG TTTTAAACCC GCTGCAAAAA
451  ATATTTTCCC TCGTCGGCAA AAAAACCTTC CATTTCGGCG ATGTCGGCAA
501  AGGTTCGGGC GCGAAACTCG TCTTGAACTC GCTCTTGGGC ATTTTCGGCG
551  AAGCGTACAG CGAAGCGATG CTGATGGCGC GGCAGTTCGG CATCGATACC
601  GACACCATCG TCGAAGCCAT CGGCGGCTCG GCAATGGACT CGCCCATGTT
651  CCAAACCAAA AAATCCCTGT GGGCAAACCG CGAATTCCCA CCCGCCTTCG
701  CCCTCAAACA CGCCTCCAAA GACCTCAACC TCGCCGTCAA AGAGCTTGAA
751  CAGGCAGGCA ACACCCTGCC CGCCGTCGAA ACCGTTGCTG CCAGCTACCG
801  CAAAGCAGTC GAAGCCGGCT ACGGCGAACA GGACGTTTCC GGCGTTTACC
851  TGAAATTGGC AGAACACTGA
```

This corresponds to the amino acid sequence <SEQ ID 374; ORF 105-1.a>:

a105-1.pep

```
  1  MSANEYTQIG WIGLGQMGLP MVTRLLDGGI EVGVYNRSPD KTAPISAKGA
 51  KVYGNTAELV RDYPVIFLMV SDYAAVCDIL NGVRDGLAGK IIVNMSTISP
101  TENLAVKALV EAAGGQFAEA PVSGSVGPAT NGTLLILFGG SEAVLNPLQK
151  IFSLVGKKTF HFGDVGKGSG AKLVLNSLLG IFGEAYSEAM LMARQFGIDT
201  DTIVEAIGGS AMDSPMFQTK KSLWANREFP PAFALKHASK DLNLAVKELE
251  QAGNTLPAVE TVAASYRKAV EAGYGEQDVS GVYLKLAEH*
```

a105-1/m105-1 99.0% identity in 289 aa overlap

```
                    10        20        30        40        50        60
a105-1.pep  MSANEYTQIGWIGLGQMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGNTAELV
            ||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
m105-1       MSANEYAQIGWIGLGQMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGNTAELV
                     10        20        30        40        50        60

                     70        80        90       100       110       120
a105-1.pep   RDYPVIFLMVSDYAAVCDILNGVRDGLAGKIIVNMSTISPTENLAVKALVEAAGGQFAEA
             |||||||||||||||||||||||||||||||||||||||||||||||||' |||||||||
m105-1       RDYPVIFLMVSDYAAVCDILNGVRDGLAGKIIVNMSTISPTENLAVKALVEAAGGQFAEA
                     70        80        90       100       110       120

                    130       140       150       160       170       180
a105-1.pep   PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFHFGDVGKGSGAKLVLNSLLG
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m105-1       PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFHFGDVGKGSGAKLVLNSLLG
                    130       140       150       160       170       180

                    190       200       210       220       230       240
a105-1.pep   IFGEAYSEAMLMARQFGIDTDTIVEAIGGSAMDSPMFQTKKSLWANREFPPAFALKHASK
             |||||||| |||||||||||||||||||||| ||||||||||||||||||||||||||||
m105-1       IFGEAYSEXMLMARQFGIDTDTIVEAIGDSAMDSPMFQTKKSLWANREFPPAFALKHASK
                    190       200       210       220       230       240

                    250       260       270       280       290
a105-1.pep   DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAEHX
             |||||||||||||||||'||||||||||||||||||||||||||||||||
m105-1       DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAEHX
                    250       260       270       280       290
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 375>:

```
g107.seq
      1  ATGGTATTAA CCTTTATTTG GGCAACCGGC GGCCTGGTTG CCGATGCCAA
     51  ACCGTCCGTC CTCTTCGACA CCCAAGCCCC CGTCGGCACC GGCTACTGGA
    101  TTTACGCCGC CACCGCCCTG CCCGTCTGCC TCGCTTCCTT CGGCTTCCAC
    151  GGCAACGTTT CCAGCCTGCT CAAATACTTT AAAGGCGACG cgcCcaaagt
    201  GgCGAaATCc atctggGcag gtacattggT TGCCttggta atttacgtcc
    251  TCTggcaaac cgccatCcaa agcaaccTGC cgcgcaacga gttcgcCCCc
    301  gtgattgccg ccgagaggca actCTCCGTC CTgaatgaaa cccTGtccaa
    351  attcgcccaa accggcgata tggataAaat attgtcccta tttccctaca
    401  tggcaatcgc cacctccttt ttaggcgTAA Ccttaggcct gtttgacaac
    451  atcgccggac atcttcaaat ggaacgacag tatgtccggg cggcaccaaa
    501  accgtcgcgc tga
```

This corresponds to the amino acid sequence <SEQ ID 376; ORF 107.ng>:

```
g107.pep
      1  MVLTFIWATG GLVADAKPSV LFDTQAPVGT GYWIYAATAL PVCLASFGFH
     51  GNVSSLLKYF KGDAPKVAKS IWAGTLVALV IYVLWQTAIQ SNLPRNEFAP
    101  VIAAERQLSV LNETLSKFAQ TGDMDKILSL FPYMAIATSF LGVTLGLFDN
    151  IAGHLQMERQ YVRAAPKPSR *
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 377>:

```
m107.seq
      1  ATGGTATTGA CCTTTATTTG GGCGGCCGGC GGGCTGATTG CCGATGCCAA
     51  GCCGTCCGTC CTCTTCGATA CCCAAGCCCC CGCCGGCACA AACTACTGGA
    101  TTTACGCCGs CACCGCCCTG CCCGTCTGCC TCGCTTCCTT CGGCTTCCAC
    151  GGCAACGTCT CCAGCCTGCT CAAATACTTT AAAGGCGACG CGCCCAAAGT
    201  GGCTAAATCC ATCTGGACGG GCACACTGAT TGCGCTGGTA ATTTACGTCC
    251  TCTGGCAAAC CGCCATCCAA GGCAACCTGC CGCGCAACGA GTTCGCCCCC
    301  GTCATCGCCG CCGAAGGGCA AGTCTCCGTC CTCATCGAAA CCCTGTCCAA
    351  ATTCGCCCAA ACCGGCAATA TGGACAAAAT ATTGTCCCTG TTTTCCTATA
    401  TGGCGATCGC CACCTCGTTT TTAGGCGTAA CGCTCGGACT CTTCGACTAC
    451  ATCGCCCATC TTCAAATGGA ACGACAGCAT CTCCGGgCCG CACCAAAACC
    501  GCCGCGCTGA
```

This corresponds to the amino acid sequence <SEQ ID 378; ORF 107>:

```
m107.pep..
      1  MVLTFIWAAG GLIADAKPSV LFDTQAPAGT NYWIYAXTAL PVCLASFGFH
     51  GNVSSLLKYF KGDAPKVAKS IWTGTLIALV IYVLWQTAIQ GNLPRNEFAP
```

```
101  VIAAEGQVSV LIETLSKFAQ TGNMDKILSL FSYMAIATSF LGVTLGLFDY
151  IAHLQMERQH LRAAPKPPR*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 107 shows 89.4% identity over a 170 aa overlap with a predicted ORF (ORF 107.ng) from *N. gonorrhoeae*:

m107/g107

```
                       10        20        30        40        50        60
m107.pep    MVLTFIWAAGGLIADAKPSVLFDTQAPAGTNYWIYAXTALPVCLASFGFHGNVSSLLKYF
            ||||||||:|||:||||||||||||||:|| |||| ||||||||||||||||||||||||
g107        MVLTFIWATGGLVADAKPSVLFDTQAPVGTGYWIYAATALPVCLASFGFHGNVSSLLKYF
                       10        20        30        40        50        60

                       70        80        90       100       110       120
m107.pep    KGDAPKVAKSIWTGTLIALVIYVLWQTAIQGNLPRNEFAPVIAAEGQVSVLIETLSKFAQ
            ||||||||||||:|||:|||||||||||||:|||||||||||||| |:||| ||||||||
g107        KGDAPKVAKSIWAGTLVALVIYVLWQTAIQSNLPRNEFAPVIAAERQLSVLNETLSKFAQ
                       70        80        90       100       110       120

                      130       140       150       160       170
m107.pep    TGNMDKILSLFSYMAIATSFLGVTLGLFDYIA-HLQMERQHLRAAPKPPR
            ||:|||||||| ||||||||||||||||| || |||||||::|||||| |
g107        TGDMDKILSLFPYMAIATSFLGVTLGLFDNIAGHLQMERQYVRAAPKPSR
                      130       140       150       160       170
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 379>:

a107.seq

```
  1  ATGGTATTAA CCTTTATTTG GGCAACCGGC GGCCTGATTG CCGATGCCAA
 51  ACTGCCCGTC CTCTTCGACA CCCAAGCCCC TACCGGCACC AACTACTGGA
101  TTTATGTCGC CACCGCCCTG CCCGTCTGCC TTGCGTCATT CGGTTTCCAC
151  GGCAACGTCT CCAGCCTGCT CAAATACTTT AAAGGCGACG CGCCCAAAGT
201  GGCTAAATCC ATCTGGACGG GCACACTGAT TGCGCTGGTA ATTTACGTCC
251  TCTGGCAAAC CGCCATCCAA GGCAACCTGC CGCGCAACGA GTTCGCCCCC
301  GTGATTGCCG CCGAAGGGCA AGTCTCCGTC CTGATTGAAA CCCTGTCCAA
351  ATTCGCCCAA ACCGGCAATA TGGACAAAAT ATTGTCCCTG TTTTCCTATA
401  TGGCGATCGC CACCTCGTTT TTAGGCGTAA CGCTCGGACT CTTCGACTAC
451  ATCGCCGACA TCTTCAAATG GAACGACAGC GTGTCCGGCC GCACCAAAAC
501  CGCCGCGCTG ACCTTCCTGC CGCCTCTAAT TTCCTGCCTG CTCTTCCCCA
551  CCGGCTTTGT TACCGCCATC GGCTACGTCG GCCTGGCGGC AACCGTCTGG
601  ACAGGCATCA TCCCCGCCAT GCTGCTCTAC CGTTCGCGCA AAAAATTCGG
651  CGCAGGCAAA ACCTATAAAG TTTACGGCGG CTTGTGGCTG ATGGTTTGGG
701  TCTTCCTTTT CGGCATCGTC AACATCGCCG CACAGGTATT GAGCCAAATG
751  GAACTCGTCC CCGTATTTAA AGGATAA
```

This corresponds to the amino acid sequence <SEQ ID 380; ORF 107.a>:

a107.pep

```
  1  MVLTFIWATG GLIADAKLPV LFDTQAPTGT NYWIYVATAL PVCLASFGFH
 51  GNVSSLLKYF KGDAPKVAKS IWTGTLIALV IYVLWQTAIQ GNLPRNEFAP
101  VIAAEGQVSV LIETLSKFAQ TGNMDKILSL FSYMAIATSF LGVTLGLFDY
151  IADIFKWNDS VSGRTKTAAL TFLPPLISCL LFPTGFVTAI GYVGLAATVW
201  TGIIPAMLLY RSRKKFGAGK TYKVYGGLWL MVWVFLFGIV NIAAQVLSQM
251  ELVPVFKG*
```

m107/a107   94.8% identity in 154 aa overlap

```
                       10        20        30        40        50        60
m107.pep    MVLTFIWAAGGLIADAKPSVLFDTQAPAGTNYWIYAXTALPVCLASFGFHGNVSSLLKYF
            ||||||||:||||||||  ||||||||:|||||||: |||||||||||||||||||||||
a107        MVLTFIWATGGLIADAKLPVLFDTQAPTGTNYWIYVATALPVCLASFGFHGNVSSLLKYF
                       10        20        30        40        50        60
```

```
                70        80        90       100       110       120
m107.pep   KGDAPKVAKSIWTGTLIALVIYVLWQTAIQGNLPRNEFAPVIAAEGQVSVLIETLSKFAQ
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a107       KGDAPKVAKSIWTGTLIALVIYVLWQTAIQGNLPRNEFAPVIAAEGQVSVLIETLSKFAQ
                70        80        90       100       110       120

               130       140       150       160       170
m107.pep   TGNMDKILSLFSYMAIATSFLGVTLGLFDYIAHLQMERQHLRAAPKPPRX
           |||||||||||||||||||||||||||||||| :
a107       TGNMDKILSLFSYMAIATSFLGVTLGLFDYIADIFKWNDSVSGRTKTAALTFLPPLISCL
               130       140       150       160       170       180

a107       LFPTGFVTAIGYVGLAATVWTGIIPAMLLYRSRKKFGAGKTYKVYGGLWLMVWVFLFGIV
               190       200       210       220       230       240
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 381>:

```
g108.seq
      1  ATGttgccgg gCTTCAACCG GATATTCAaa cggTTTGCTC CAACACTCGG
     51  AAcggCGCAT AAAACGCCgc ccTTCGCGTT ATCCCGAACG GGGCGCCTAA
    101  TCAGATCCTA TCGCCATAAA AGGCGGGGTT TCAACCGAAA AGGAATTGAG
    151  ATGAATAAAA CCTTGTCTAT TTTGCCGGCG GCAATCTTAC TCGGCGGGTG
    201  CGCCGCCGGC GGCAACACAT TCGGCAGCTT AGACGGCGGC ACGGGTATGG
    251  GTGGCAGCAT CGTCAAAATG ACGGTAGAAA gccAATGCCG TGCGGAATTG
    301  GACAGGCGCA GCGAATGGCG TTTGACCGCG CTGGCGATGA GTGCCGAAAA
    351  ACAGGCGGAA TGGGAAAACA AGATTTGCGG CTGCGCTACC GAAGAAGCAC
    401  CTAACCAGCT GACCGGCAAC GATGTGATGC AGATGCTGAa ccagtccacG
    451  CGCaatcagg cacTtgccgc CCtgaccgTC AAAacggtTT CcgcctgcTT
    501  CAaacgcctg tACCGCTAa
```

This corresponds to the amino acid sequence <SEQ ID 382; ORF 108.ng>:

```
g108.pep
      1  MLPGFNRIFK RFAPTLGTAH KTPPFALSRT GRLIRSYRHK RRGFNRKGIE
     51  MNKTLSILPA AILLGGCAAG GNTFGSLDGG TGMGGSIVKM TVESQCRAEL
    101  DRRSEWRLTA LAMSAEKQAE WENKICGCAT EEAPNQLTGN DVMQMLNQST
    151  RNQALAALTV KTVSACFKRL YR*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 383>:

```
m108.seq
      1  ATGTTGCCGG GCTTCAACCG GATATTCAAA CGGTTTGTTC CAACACTCGG
     51  AACGGCGCAT AAAACGCCGC CCTTCGCGTT ATCCCGAACG GGGCGGCTAA
    101  TCAGATTCTA TCGCCATAAA AGGCGGGGTT TCAACCGAAA AGGAATTGAG
    151  ATGAATAAAA CCTTGTCTAT TTTGCCGGTG GCAATCTTAC TCGGCGGCTG
    201  CGCCGCCGGA GGCGGTAACA CATTCGGCAG CTTAGACGGT GGCACAGGCA
    251  TGGGCGGCAG CATCGTCAAA ATGGCGGTTG GGAGCCAATG CCGTGCGGAA
    301  TTGGACAAAC GCAGCGAATG GCGTTTGACC GCGCTGGCGA TGAGTGCCGA
    351  AAAACAGGCG GAGTGGGAAA ACAAGATTTG CGCTTGCGTC GCCCAAGAAG
    401  CACCCGAACG GATGACCGGC AACGATGTGA TGCAGATGCT GGCTCCGTCC
    451  ACGCGCAATC AGGCACTTGC CGCCCTGACC GCCAAAACGG TTTCCGCCTG
    501  CTTCAAACAC CTGTACCGCT AA
```

This corresponds to the amino acid sequence <SEQ ID 384; ORF 108>:

```
m108.pep
      1  MLPGFNRIFK RFVPTLGTAH KTPPFALSRT GRLIRFYRHK RRGFNRKGIE
     51  MNKTLSILPV AILLGGCAAG GGNTFGSLDG GTGMGGSIVK MAVGSQCRAE
    101  LDKRSEWRLT ALAMSAEKQA EWENKICACV AQEAPERMTG NDVMQMLAPS
    151  TRNQALAALT AKTVSACFKH LYR*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 108 shows 89.6% identity over a 173 aa overlap with a predicted ORF (ORF 108.ng) from *N. gonorrhoeae*:

```
m108/g108
```

```
                   10        20        30        40        50        60
m108.pep   MLPGFNRIFKRFVPTLGTAHKTPPFALSRTGRLIRFYRHKRRGFNRKGIEMNKTLSILPV
           |||||||||||:||||||||||||||||||||||| |||||||||||||||||||||||:
g108       MLPGFNRIFKRFAPTLGTAHKTPPFALSRTGRLIRSYRHKRRGFNRKGIEMNKTLSILPA
                   10        20        30        40        50        60

                   70        80        90       100       110       120
m108.pep   AILLGGCAAGGGNTFGSLDGGTGMGGSIVKMAVGSQCRAELDKRSEWRLTALAMSAEKQA
           ||||||||||| |||||||||||||||||||:| ||||||||:|||||||||||||||||
g108       AILLGGCAAGG-NTFGSLDGGTGMGGSIVKMTVESQCRAELDRRSEWRLTALAMSAEKQA
                   70        80        90       100       110

                  130       140       150       160       170
m108.pep   EWENKICACVAQEAPERMTGNDVMQMLAPSTRNQALAALTAKTVSACFKHLYRX
           |||||||:|::::|||:::|||||||||  |||||||||||:||||||||:|||
g108       EWENKICGCATEEAPNQLTGNDVMQMLNQSTRNQALAALTVKTVSACFKRLYRX
         120       130       140       150       160       170
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 385>:

a108.seq

```
  1  ATGTTGCCGG GCTTCAACCG GATATTCAAA CGGTTTGTTC CAACACTCGG
 51  AACGGCGCAT AAAACGCCGC CCTTCGCGTT ATCCCGAACG GGCGGCTAA
101  TCAGATTCTA TCGCCATAAA AGGCGGGGTT TCAACCGAAA AGGAATTGAG
151  ATGAATAAAA CCTTGTCTAT TTTGCCGGTG GCAATCTTAC TCGGCGGCTG
201  CGCCGCCGGG GGCGGTAACA CATTCGGCAG CTTAGACGGC GGCACAGGTA
251  TGGGCGGCAG CATCGTCAAA ATGGCGGTAG AAAGCCAATG CCGTGCGGAA
301  TTGAACAAAC GCAGCGAATG GCGTTTGACC GCGCTGGCGA TGAGTGCCGA
351  AAAACAGGCG GAATGGGAAA ACAAGATTTG CGCTTGCGTC GCCCAAGAAG
401  CACCCAACCA GCTGACCGGC AACGATGTGA TGCAGATGCT GGATCCGTCC
451  ACGCGCAATC AGGCACTTGC CGCCCTGACC GCCAAAACGG TTTCCGCCTG
501  CTTCAAACAC CTGTACCGCT AA
```

This corresponds to the amino acid sequence <SEQ ID 386; ORF 108.a>:

a108.pep

```
  1  MLPGFNRIFK RFVPTLGTAH KTPPFALSRT GRLIRFYRHK RRGFNRKGIE
 51  MNKTLSILPV AILLGGCAAG GGNTFGSLDG GTGMGGSIVK MAVESQCRAE
101  LNKRSEWRLT ALAMSAEKQA EWENKICACV AQEAPNQLTG NDVMQMLDPS
151  TRNQALAALT AKTVSACFKH LYR*
```

m108/a108 96.5% identity in 173 aa overlap

```
                   10        20        30        40        50        60
m108.pep   MLPGFNRIFKRFVPTLGTAHKTPPFALSRTGRLIRFYRHKRRGFNRKGIEMNKTLSILPV
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a108       MLPGFNRIFKRFVPTLGTAHKTPPFALSRTGRLIRFYRHKRRGFNRKGIEMNKTLSILPV
                   10        20        30        40        50        60

                   70        80        90       100       110       120
m108.pep   AILLGGCAAGGGNTFGSLDGGTGMGGSIVKMAVGSQCRAELDKRSEWRLTALAMSAEKQA
           ||||||||||||||||||||||||||||||||| ||||||||:|||||||||||||||||
a108       AILLGGCAAGGGNTFGSLDGGTGMGGSIVKMAVESQCRAELNKRSEWRLTALAMSAEKQA
                   70        80        90       100       110       120

                  130       140       150       160       170
m108.pep   EWENKICACVAQEAPERMTGNDVMQMLAPSTRNQALAALTAKTVSACFKHLYRX
           |||||||||||||||:::||||||||| ||||||||||||||||||||||||||
a108       EWENKICACVAQEAPNQLTGNDVMQMLDPSTRNQALAALTAKTVSACFKHLYRX
                  130       140       150       160       170
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 387>:

```
g109.seq
      1  ATGTATTATC GCCGGGTTGT GGGGCTATCC GATGGACTTG GCGATTTGGC
     51  AGCCGGTATT GATCGTAGGC GTATGCTTAC CGCTTTTGGA AGCGGGCATG
    101  GAAATGACGC GCAAAGGCAA AACCACCCAA TCCGCCGCCA TCGTGGTGTT
    151  CTCTTCCGTC TGGTCAATCC GGTTTTCGGC TGGGCGTTGA CGATGCTGTT
    201  GGATAATTTG GGCTTAATCG GCTGCAAAGA ACGCAGCGCG CAATTAGGTT
    251  TTGTCGGACG AGTATTGATA CCCGCAGTAG GTTTCTTAAT CTTGTGTGTG
    301  GCGATGGGTG CGGTCGGGAT GCTGCCCGGT ATCCCTCCGT TTTTGGAGCA
    351  GTTCAAATCT TTGGGCTAG
```

This corresponds to the amino acid sequence <SEQ ID 388; ORF 109.ng>:

```
g109.pep
      1  MYYRRVVGLS DGLGDLAAGI DRRRMLTAFG SGHGNDAQRQ NHPIRRHRGV
     51  LFRLVNPVFG WALTMLLDNL GLIGCKERSA QLGFVGRVLI PAVGFLILCV
    101  AMGAVGMLPG IPPFLEQFKS LG
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 389>:

```
m109.seq
      1  ATGTATTATC GCCGGGTTAT GGGGCTATCC GATGGACTTG GCGATTTGGC
     51  AGCCGGTATT GAGCGTAGCC TTGGTCGTAG GCGTATACTT ACCGCTTTTG
    101  GAAGCGGGCA TGGAAATGAC GCGCAAAGGC AAAACCACCC AATCCGCCGC
    151  CATCGTGGTG TTCTCTTCCG CCTTGTCAAT CCGGTTTTCG GCTGGGCGTT
    201  GACGATGCTG TTGGATAATT TGGGCTTAAT CGGCTGCAAA GAGCGCAGTG
    251  CGCAATTAGG TTTCGCCGGA CGCGTGTTGA TACCCGCAGT AGGTTTCTTG
    301  ATCTTGTGTG TGGCGATGGG TGCGGTCGGG ATGCTGCCCG GTATCCCGCC
    351  GTTTTTGGAA CACTTCAAAT CTTTGGGCTA G
```

This corresponds to the amino acid sequence <SEQ ID 4; ORF 109>:

```
m109.pep
      1  MYYRRVMGLS DGLGDLAAGI ERSLGRRRIL TAFGSGHGND AQRQNHPIRR
     51  HRGVLFRLVN PVFGWALTML LDNLGLIGCK ERSAQLGFAG RVLIPAVGFL
    101  ILCVAMGAVG MLPGIPPFLE HFKSLG*
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

ORF 109 shows 92.9% identity over a 126 aa overlap with a predicted ORF (ORF 109.ng) from *N. gonorrhoeae:*

```
m109/g109

                    10        20        30        40        50        60
m109.pep    MYYRRVMGLSDGLGDLAAGIERSLGRRRILTAFGSGHGNDAQRQNHPIRRHRGVLFRLVN
            ||||||:|||||||||||||:|    ||:||||||||||||||||||||||||||||||||
g109        MYYRRVVGLSDGLGDLAAGIDR----RRMLTAFGSGHGNDAQRQNHPIRRHRGVLFRLVN
                    10        20            30        40        50

                    70        80        90       100       110       120
m109.pep    PVFGWALTMLLDNLGLIGCKERSAQLGFAGRVLIPAVGFLILCVAMGAVGMLPGIPPFLE
            ||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
g109        PVFGWALTMLLDNLGLIGCKERSAQLGFVGRVLIPAVGFLILCVAMGAVGMLPGIPPFLE
              60        70        80        90       100       110


m109.pep    HFKSLGX
            :|||||
g109        QFKSLGX
             120
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 391>:

```
a109.seq
      1  ATGTATTATC GCCGGGTTGT GGGGCTATCC GATGGACTTG GCGATTTGGC
     51  AGCCGGTATT GAGCGTAGCC TTGGTCGTAG GCGTATACTT ACCGCTTTTG
    101  GAAGCGGGCA TGGAAATGAC GCGCAAAGGC AAAACCACCC AATCCGCCGC
    151  CACCGTGGTG TTCTCTTCCG CTTGGTCAAT CCGGTTTTCG GCTGGGCGTT
```

```
201  GACGATGCTG TTGGATAATT TGGGCTTAAT CGGCTGCAAA GAGCGCAGCG
251  CGCAATTAGG TTTCACCGGA CGCGTATTGA TACCCGTAGT AGGTTTCTTG
301  ATCTTGTGTG TGGCGATGGG TGCGGTCGGG ATGCTGCCCG GTATCCCGCC
351  GTTTTTGGAG CACTTCAAAT CTTTGGGCTA G
```

This corresponds to the amino acid sequence <SEQ ID 392; ORF 109>:

a109.pep

```
  1  MYYRRVVGLS DGLGDLAAGI ERSLGRRRIL TAFGSGHGND AQRQNHPIRR
 51  HRGVLFRLVN PVFGWALTML LDNLGLIGCK ERSAQLGFTG RVLIPVVGFL
101  ILCVAMGAVG MLPGIPPFLE HFKSLG*
```

m109/a109   97.6% identity in 126 aa overlap

```
                   10        20        30        40        50        60
m109.pep    MYYRRVMGLSDGLGDLAAGIERSLGRRRILTAFGSGHGNDAQRQNHPIRRHRGVLFRLVN
            ||||||:|||||||||||||||||||||||||||||.|||||||||||||||||||||||
a109        MYYRRVVGLSDGLGDLAAGIERSLGRRRILTAFGSGHGNDAQRQNHPIRRHRGVLFRLVN
                   10        20        30        40        50        60

                   70        80        90       100       110       120
m109.pep    PVFGWALTMLLDNLGLIGCKERSAQLGFAGRVLIPAVGFLILCVAMGAVGMLPGIPPFLE
            ||||||||||||||||||||||||||||:||||||:||||||||||||||||||||||||
a109        PVFGWALTMLLDNLGLIGCKERSAQLGFTGRVLIPVVGFLILCVAMGAVGMLPGIPPFLE
                   70        80        90       100       110       120

m109.pep    HFKSLGX
            |||||||
a109        HFKSLGX
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 393>:

g111.seq

```
  1  ATGCCGTCTG AAACACGCCT GCCGAACCTT ATCCGCGCCT TGATATTTGC
 51  CCTGGGTTTC ATCTTCCTGA ACGCCTGTTC GGaacaaacC GCGCAaaccg
101  TTACCCTGCA AGGCGAAACG ATGGGTACGA CCtATACCGT CAAATACCTT
151  TCAAATAATC GGGACAAACT CCCCTCCCCT GCCAAAATAC AAAAGCGCAT
201  TGATGATGCG CTTAAAGAAG TCAACCGGCA GATGtccaCC TACCAGACCG
251  ATTCCGAAAT CAGCCGGTTt atacagacan atgctggaga gctcttcgcg
301  tntcatgcag nttctataac tgattccgcc gaagactgtc tgcctaatac
351  gcctatctca tcggcgctct ga
```

This corresponds to the amino acid sequence <SEQ ID 394; ORF 111.ng>:

g111.pep

```
  1  MPSETRLPNL IRALIFALGF IFLNACSEQT AQTVTLQGET MGTTYTVKYL
 51  SNNRDKLPSP AKIQKRIDDA LKEVNRQMST YQTDSEISRF IQTAGELFAH
101  ASITDSAEDC LPNTPISSAL *
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 395>:

m111.seq

```
  1  ATGCCGTCTG AAACACGCCT GCCGAACTTT ATCCGCGTCT TGATATTTGC
 51  CCTGGGTTTC ATCTTCCTGA ACGCCTGTTC GGAACAAACC GCGCAAACCG
101  TTACCCTGCA AGGCGAAACG ATGGGCACGA CCTATAyCGT CAAATACCTT
151  TCAAATAATC GGGACAAACT CCCCTCACCT GCCGAAATAC AwAAACGCAT
201  CGATGACGCG CTTAAAGAAk TCAACCGGyA GATGTCCACC TATCAGCCCG
251  ACTCCGAAAT CAGCCGGTTC AACCAACACA CAGCCGGCAA GCCCCTCCGC
301  ATTTCAAGCG ACTTCGCACA CGTTACTGCC GAAGCCGTCC GCCTGAACCG
351  CCTGACACAC GGCGCGCTGG ACGTAACCGT CGGCCCCTTG GTCAACCTTT
401  GGGGATTCGG CCCCGACAAA TCCGTTACCC GTGAACCGTC GCCGGAACAA
451  ATCAAACAGG CGGCATCTTA TACGGGCATA GACAAAATCA TTTTGAAACA
501  AGGCAAAGAT TACGCTTCCT TGAGCAAAAC CCACCCCAAG GCCTATTTGG
551  ATTTATCTTC GATTGCCAAA GGCTTCGGCG TTGATAAAGT TGCGGGCGAA
```

```
 601  CTGGAAAAAT ACGGCATTCA AAATTATCTG GTCGAAATCG GCGGCGAGTT
 651  GCACGGCAAA GGCAAAAACG CGCGCGGCGA ACCGTGGCGC ATCGGTATCG
 701  AGCAGCCCAA TATCGTCCAA GGCGGCAATA CGCAGATTAT CGTCCCGCTG
 751  AACAACCGTT CGCTTGCCAC TTCCGGCGAT TACCGTATTT TCCACGTCGA
 801  TAAAAACGGC AAACGCCTCT CCCATATCAT CAACCCGAAC AACAAACGAC
 851  CCATCAGCCA CAACCTCGCC TCCATCAGCG TGGTCGCAGA CAGTGCGATG
 901  ACGGCGGACG GCTTGTCCAC AGGATTATTC GTATTGGGCG AAACCGAAGC
 951  CTTAAAGCTG GCAGAGCGCG AAAAACTCGC TGTTTTCCTG ATTGTCAGGG
1001  ATAAAGGCGG cTACCGCACC GCCATGTCTT CCGAATTTGA AAAACTGcTC
1051  CGCTAA
```

This corresponds to the amino acid sequence <SEQ ID 396; ORF 111>:

```
m111.pep
   1  MPSETRLPNF IRVLIFALGF IFLNACSEQT AQTVTLQGET MGTTYXVKYL
  51  SNNRDKLPSP AEIXKRIDDA LKEXNRXMST YQPDSEISRF NQHTAGKPLR
 101  ISSDFAHVTA EAVRLNRLTH GALDVTVGPL VNLWGFGPDK SVTREPSPEQ
 151  IKQAASYTGI DKIILKQGKD YASLSKTHPK AYLDLSSIAK GFGVDKVAGE
 201  LEKYGIQNYL VEIGGELHGK GKNARGEPWR IGIEQPNIVQ GGNTQTIVPL
 251  NNRSLATSGD YRIFHVDKNG KRLSHIINPN NKRPISHNLA SISVVADSAM
 301  TADGLSTGLF VLGETEALKL AEREKLAVFL IVRDKGGYRT AMSSEFEKLL
 351  R*
```

ORF 111 shows 88.7% identity over a 97 aa overlap with a predicted ORF (ORF 111.ng) from *N. gonorrhoeae:*

```
m111.pep/g111.pep

                    10        20        30        40        50        60
m111.pep    MPSETRLPNFIRVLIFALGFIFLNACSEQTAQTVTLQGETMGTTYXVKYLSNNRDKLPSP
            |||||||||:||:|||||||||||||||||||||||||||||||:|||||||||||||||
g111        MPSETRLPNLIRALIFALGFIFLNACSEQTAQTVTLQGETMGTTYTVKYLSNNRDKLPSP
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m111.pep    AEIXKRIDDALKEXNRXMSTYQPDSEISRFNQHTAGKPLRISSDFAHVTAEAVRLNRLTH
            |:| ||||||||| || ||||| ||||||| |  :||:
g111        AKIQKRIDDALKEVNRQMSTYQTDSEISRFIQTXAGELFAXHAXSITDSAEDCLPNTPIS
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m111.pep    GALDVTVGPLVNLWGFGPDKSVTREPSPEQIKQAASYTGIDKIILKQGKDYASLSKTHPK

g111        SALX
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 397>:

a111.seq

```
   1  ATGCCGTCTG AAACACGCCT GCCGAACTTT ATCCGCACCT TGATATTTGC
  51  CCTGAGTTTT ATCTTCCTGA ACGCCTGTTC GGAACAAACC GCGCAAACCG
 101  TTACCCTGCA AGGTGAAACG ATGGGCACGA CCTATACCGT CAAATACCTT
 151  TCAAATAATC GGGACAAACT CCCCTCACCT GCCGAAATAC AAAAGCGCAT
 201  CGATGACGCG CTTAAAGAAG TCAACCGGCA GATGTCCACC TATCAGCCCG
 251  ACTCCGAAAT CAGCCGGTTC AACCAACACA CAGCCGGCAA GCCCCTCCGC
 301  ATTTCAAGCG ACTTCGCACA CGTTACTGCC GAAGCCGTCC ACCTGAACCG
 351  CCTGACACAC GGCGCGCTGG ACGTAACCGT CGGCCCCTTG GTCAACCTTT
 401  GGGGATTCGG CCCCGACAAA TCCGTTACCC GTGAACCGTC GCCGGAACAA
 451  ATCAAACAAG CAGCATCTTA TACGGGCATA GACAAAATCA TTTTGAAACA
 501  AGGCAAAGAT TACGCTTCCT TGAGCAAAAC CCACCCCAAG GCCTATTTGG
 551  ATTTATCTTC GATTGCCAAA GGCTTCGGCG TTGATAAAGT TGCGGGCGAA
 601  CTGGAAAAAT ACGGCATTCA AAATTATCTG GTCGAAATCG GCGGCGAGTT
 651  GCACGGCAAA GGCAAAAACG CGCGCGGCGA ACCTTGGCGC ATCGGCATCG
 701  AACAGCCCAA CATCGTCCAA GGCGGCAATA CGCAGATTAT CGTCCCGCTG
 751  AACAACCGTT CGCTTGCCAC TTCCGGCGAT TACCGTATTT TCCACGTCGA
```

```
 801  TAAAAGCGGC AAACGCCTCT CCCATATCAT TAATCCGAAC AACAAACGAC
 851  CCATCAGCCA CAACCTCGCC TCCATCAGCG TGGTCGCAGA CAGTGCGATG
 901  ACGGCGGACG GCTTGTCCAC AGGATTATTC GTATTGGGCG AAACCGAAGC
 951  CTTAAAGCTG GCAGAGCGCG AAAAACTCGC TGTTTTCCTG ATTGTCAGGG
1001  ATAAAGGCGG CTACCGCACC GCCATGTCTT CCGAATTTGA AAAACTGCTC
1051  CGCTAA
```

This corresponds to the amino acid sequence <SEQ ID 398; ORF 111.a>:

a111.pep

```
  1  MPSETRLPNF IRTLIFALSF IFLNACSEQT AQTVTLQGET MGTTYTVKYL
 51  SNNRDKLPSP AEIQKRIDDA LKEVNRQMST YQPDSEISRF NQHTAGKPLR
101  ISSDFAHVTA EAVHLNRLTH GALDVTVGPL VNLWGFGPDK SVTREPSPEQ
151  IKQAASYTGI DKIILKQGKD YASLSKTHPK AYLDLSSIAK GFGVDKVAGE
201  LEKYGIQNYL VEIGGELHGK GKNARGEPWR IGIEQPNIVQ GGNTQIIVPL
251  NNRSLATSGD YRIFHVDKSG KRLSHIINPN NKRPISHNLA SISVVADSAM
301  TADGLSTGLF VLGETEALKL AEREKLAVFL IVRDKGGYRT AMSSEFEKLL
351  R*
```

m111/a111 97.7% identity in 351 aa overlap

```
                        10        20        30        40        50        60
m111.pep       MPSETRLPNFIRVLIFALGFIFLNACSEQTAQTVTLQGETMGTTYXVKYLSNNRDKLPSP
               ||||||||||||:|||||:|||||||||||||||||||||||||:|||||||||||||||
a111           MPSETRLPNFIRTLIFALSFIFLNACSEQTAQTVTLQGETMGTTYTVKYLSNNRDKLPSP
                        10        20        30        40        50        60

                        70        80        90       100       110       120
m111.pep       AEIXKRIDDALKEXNRXMSTYQPDSEISRFNQHTAGKPLRISSDFAHVTAEAVRLNRLTH
               ||| ||||||||| || ||||||||||||||||||||||||||||||||||||:||||||
a111           AEIQKRIDDALKEVNRQMSTYQPDSEISRFNQHTAGKPLRISSDFAHVTAEAVHLNRLTH
                        70        80        90       100       110       120

                       130       140       150       160       170       180
m111.pep       GALDVTVGPLVNLWGFGPDKSVTREPSPEQIKQAASYTGIDKIILKQGKDYASLSKTHPK
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a111           GALDVTVGPLVNLWGFGPDKSVTREPSPEQIKQAASYTGIDKIILKQGKDYASLSKTHPK
                       130       140       150       160       170       180

                       190       200       210       220       230       240
m111.pep       AYLDLSSIAKGFGVDKVAGELEKYGIQNYLVEIGGELHGKGKNARGEPWRIGIEQPNIVQ
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a111           AYLDLSSIAKGFGVDKVAGELEKYGIQNYLVEIGGELHGKGKNARGEPWRIGIEQPNIVQ
                       190       200       210       220       230       240

                       250       260       270       280       290       300
m111.pep       GGNTQIIVPLNNRSLATSGDYRIFHVDKNGKRLSHIINPNNKRPISHNLASISVVADSAM
               |||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
a111           GGNTQIIVPLNNRSLATSGDYRIFHVDKSGKRLSHIINPNNKRPISHNLASISVVADSAM
                       250       260       270       280       290       300

                       310       320       330       340       350
m111.pep       TADGLSTGLFVLGETEALKLAEREKLAVFLIVRDKGGYRTAMSSEFEKLLRX
               ||||||||||||||||||||||||||||||||||||||||||||||||||||
a111           TADGLSTGLFVLGETEALKLAEREKLAVFLIVRDKGGYRTAMSSEFEKLLRX
                       310       320       330       340       350
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 399>:

g111-1.seq

```
  1  ATGCCGTCTG AAACACGCCT GCCGAACCTT ATCCGCGCCT TGATATTTGC
 51  CCTGGGTTTC ATCTTCCTGA ACGCCTGTTC GGAACAAACC GCGCAAacCG
101  TTACCCTGCA AGGCGAAACG ATGGGTACGA CCTATACCGT CAAATACCTT
151  TCAAATAATC GGGACAAACT CCCCTCCCCT GCCAAAATAC AAAAGCGCAT
201  TGATGATGCG CTTAAAGAAG TCAACCGGCA GATGTCCACC TACCAGACCG
```

```
 251  ATTCCGAAAT CAGCCGGTTC AACCAACACA CAGCCGGCAA GCCCCTCCGC
 301  ATTTCAAGCG ATTTCGCACA CGTTACCGCC GAAGCCGTCC GCCTGAACCG
 351  CCTGACTCAC GGCGCACTGG ACGTAACCGT CGGCCCTTTG GTCAACCTTT
 401  GGGGGTTCGG CCCCGACAAA TCCGTTACCC GTGAACCGTC GCCGGAACAA
 451  ATCAAACAGG CGGCATCTTA TACGGGCATA GACAAAATCA TTTTGCAACA
 501  AGGCAAAGAT TACGCTTCCT TGAGCAAAAC CCACCCCAAA GCCTATTTGG
 551  ATTTATCTTC GATTGCCAAA GGCTTCGGCG TTGATAAAGT TGCGGGCGAA
 601  CTGGAAAAAT ACGGCATTCA AAATTATCTG GTCGAAATCG GCggcGAGTT
 651  GCACGGCAAA GGCAAAAATG CGCACGGCGA ACCGTGGCGC ATCGGTATAG
 701  AGCAACCCAA TATcatccaa ggcggcaata cgcAGattat cgtcccgctg
 751  aaCaaccgtt cgcttgccac ttccggcgAT taccgtaTTT tccacgtcgA
 801  TAAAAACGGC Aaacgccttt cccacATCAT CAATCCCAAC AACAAACGAC
 851  CCATCAGcCA CAAcctcgcc tcCATCAgCg TGGTCTCAGA CACTGCAATG
 901  ACGGCGGACG GTTTATCCAC AGGATTATTT GTTTTAGGCG AAACCGAAGC
 951  CTTAAGGCTG GCAGAACAAG AAAAACTCGC TGTTTTCCTA ATTGTCCGGG
1001  ATAAGGACGG CTACCGCACC GCCATGTCTT CCGAATTTGC CAAGCTGCTC
1051  CGCTAA
```

This corresponds to the amino acid sequence <SEQ ID 400; ORF 111-1.ng>:

g111-1.pep

```
   1  MPSETRLPNL IRALIFALGF IFLNACSEQT AQTVTLQGET MGTTYTVKYL
  51  SNNRDKLPSP AKIQKRIDDA LKEVNRQMST YQTDSEISRF NQHTAGKPLR
 101  ISSDFAHVTA EAVRLNRLTH GALDVTVGPL VNLWGFGPDK SVTREPSPEQ
 151  IKQAASYTGI DKIILQQGKD YASLSKTHPK AYLDLSSIAK GFGVDKVAGE
 201  LEKYGIQNYL VEIGGELHGK GKNAHGEPWR IGIEQPNIIQ GGNTQIIVPL
 251  NNRSLATSGD YRIFHVDKNG KRLSHIINPN NKRPISHNLA SISVVSDSAM
 301  TADGLSTGLF VLGETEALRL AEQEKLAVFL IVRDKDGYRT AMSSEFAKLL
 351  R*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 401>:

m111-1.seq

```
   1  ATGCCGTCTG AAACACGCCT GCCGAACTTT ATCCGCGTCT TGATATTTGC
  51  CCTGGGTTTC ATCTTCCTGA ACGCCTGTTC GGAACAAACC GCGCAAACCG
 101  TTACCCTGCA AGGCGAAACG ATGGGCACGA CCTATACCGT CAAATACCTT
 151  TCAAATAATC GGGACAAACT CCCCTCACCT GCCGAAATAC AAAAACGCAT
 201  CGATGACGCG CTTAAAGAAG TCAACCGGCA GATGTCCACC TATCAGCCCG
 251  ACTCCGAAAT CAGCCGGTTC AACCAACACA CAGCCGGCAA GCCCCTCCGC
 301  ATTTCAAGCG ACTTCGCACA CGTTACTGCC GAAGCCGTCC GCCTGAACCG
 351  CCTGACACAC GGCGCGCTGG ACGTAACCGT CGGCCCCTTG GTCAACCTTT
 401  GGGGATTCGG CCCCGACAAA TCCGTTACCC GTGAACCGTC GCCGGAACAA
 451  ATCAAACAGG CGGCATCTTA TACGGGCATA GACAAAATCA TTTTGAAACA
 501  AGGCAAAGAT TACGCTTCCT TGAGCAAAAC CCACCCCAAG GCCTATTTGG
 551  ATTTATCTTC GATTGCCAAA GGCTTCGGCG TTGATAAAGT TGCGGGCGAA
 601  CTGGAAAAAT ACGGCATTCA AAATTATCTG GTCGAAATCG GCGGCGAGTT
 651  GCACGGCAAA GGCAAAAACG CGCGCGGCGA ACCGTGGCGC ATCGGTATCG
 701  AGCAGCCCAA TATCGTCCAA GGCGGCAATA CGCAGATTAT CGTCCCGCTG
 751  AACAACCGTT CGCTTGCCAC TTCCGGCGAT TACCGTATTT TCCACGTCGA
 801  TAAAAACGGC AAACGCCTCT CCCATATCAT CAACCCGAAC AACAAACGAC
 851  CCATCAGCCA CAACCTCGCC TCCATCAGCG TGGTCGCAGA CAGTGCGATG
 901  ACGGCGGACG GCTTGTCCAC AGGATTATTC GTATTGGGCG AAACCGAAGC
 951  CTTAAAGCTG GCAGAGCGCG AAAAACTCGC TGTTTTCCTG ATTGTCAGGG
1001  ATAAAGGCGG CTACCGCACC GCCATGTCTT CCGAATTTGA AAAACTGCTC
1051  CGCTAA
```

This corresponds to the amino acid sequence <SEQ ID 402; ORF 111-1>:

m111-1.pep

```
   1  MPSETRLPNF IRVLIFALGF IFLNACSEQT AQTVTLQGET MGTTYTVKYL
  51  SNNRDKLPSP AEIQKRIDDA LKEVNRQMST YQPDSEISRF NQHTAGKPLR
 101  ISSDFAHVTA EAVRLNRLTH GALDVTVGPL VNLWGFGPDK SVTREPSPEQ
 151  IKQAASYTGI DKIILKQGKD YASLSKTHPK AYLDLSSIAK GFGVDKVAGE
 201  LEKYGIQNYL VEIGGELHGK GKNARGEPWR IGIEQPNIVQ GGNTQIIVPL
 251  NNRSLATSGD YRIFHVDKNG KRLSHIINPN NKRPISHNLA SISVVADSAM
 301  TADGLSTGLF VLGETEALKL AEREKLAVFL IVRDKGGYRT AMSSEFEKLL
 351  R*
```

m111-1/g111-1 96.6% identity in 351 aa overlap

```
                    10        20        30        40        50        60
m111-1.pep  MPSETRLPNFIRVLIFALGFIFLNACSEQTAQTVTLQGETMGTTYTVKYLSNNRDKLPSP
            ||||||||:||:||||||||||||||||||||||||||||||||||||||||||||||||
g111-1      MPSETRLPNLIRALIFALGFIFLNACSEQTAQTVTLQGETMGTTYTVKYLSNNRDKLPSP
```

```
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m111-1.pep  AEIQKRIDDALKEVNRQMSTYQPDSEISRFNQHTAGKPLRISSDFAHVTAEAVRLNRLTH
            |:|||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
g111-1      AKIQKRIDDALKEVNRQMSTYQTDSEISRFNQHTAGKPLRISSDFAHVTAEAVRLNRLTH
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m111-1.pep  GALDVTVGPLVNLWGFGPDKSVTREPSPEQIKQAASYTGIDKIILKQGKDYASLSKTHPK
            |||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
g111-1      GALDVTVGPLVNLWGFGPDKSVTREPSPEQIKQAASYTGIDKIILQQGKDYASLSKTHPK
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m111-1.pep  AYLDLSSIAKGFGVDKVAGELEKYGIQNYLVEIGGELHGKGKNARGEPWRIGIEQPNIVQ
            |||||||||||||||||||||||||||||||||||||||||||:||||||||||||||:|
g111-1      AYLDLSSIAKGFGVDKVAGELEKYGIQNYLVEIGGELHGKGKNAHGEPWRIGIEQPNIIQ
                   190       200       210       220       230       240

                   250       260       270       280       290       300
m111-1.pep  GGNTQIIVPLNNRSLATSGDYRIFHVDKNGKRLSHIINPNNKRPISHNLASISVVADSAM
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
g111-1      GGNTQIIVPLNNRSLATSGDYRIFHVDKNGKRLSHIINPNNKRPISHNLASISVVSDSAM
                   250       260       270       280       290       300

                   310       320       330       340       350
m111-1.pep  TADGLSTGLFVLGETEALKLAEREKLAVFLIVRDKGGYRTAMSSEFEKLLRX
            ||||||||||||||||||:|||:|||||||||||| |||||||||| |||||
g111-1      TADGLSTGLFVLGETEALRLAEQEKLAVFLIVRDKDGYRTAMSSEFAKLLRX
                   310       320       330       340       350
```

g111-1/p44550

```
sp|P44550|YOJL_HAEIN HYPOTHETICAL LIPOPROTEIN HI0172 PRECURSOR >gi|1074292|pir||C64144
hypothetical protein HI0172 - Haemophilus influenzae (strain Rd KW20) >gi|1573128 (U32702)
lipoprotein, putative [Haemophilus influenzae Rd] Length = 346
 Score =  349 bits (885), Expect = 2e-95
 Identities = 177/328 (53%), Positives = 240/328 (72%), Gaps = 4/328 (1%)

Query: 23  LNACSEQTAQTVTLQGETMGTTYXVKYLSNNRDKLPSPAEIXKRIDDALKEXNRXMSTYQ 82
            L AC ++T + ++L G+TMGTTY VKYL +        S  +  + I+  LK+ N  MSTY+
Sbjct: 17  LAACQKET-KVISLSGKTMGTTYHVKYLDDGSITATS-EKTHEEIEAILKDVNAKMSTYK 74

Query: 83  PDSEISRFNQHT-AGKPLRISSDFAHVTAEAVRLNRLTHGALDVTVGPLVNLWGFGPDKS 141
             DSE+SRFNQ+T    P+ IS+DFA V AEA+RLN++T GALDVTVGP+VNLWGFGP+K
Sbjct: 75  KDSELSRFNQNTQVNTPIEISADFAKVLAEAIRLNKVTEGALDVTVGPVVNLWGFGPEKR 134

Query: 142 VTREPSPEQIKQAASYTGIDKIILKQGKDYASLSKTHPKAYLDLSSIAKGFGVDKVAGEL 201
              ++P+PEQ+ +  ++ GIDKI L   K+ A+LSK  P+ Y+DLSSIAKGFGVD+VA +L
Sbjct: 135 PEKQPTPEQLAERQAWVGIDKITLDTNKEKATLSKALPQVYVDLSSIAKGFGVDQVAEKL 194

Query: 202 EKYGIQNYLVEIGGELHGKGKNARGEPWRIGIEQPNIVQGGNTQIIVPLNNRSLATSGDY 261
            E+   QNY+VEIGGE+  KGKN  G+PW+I IE+P          + ++ LNN  +A+SGDY
Sbjct: 195 EQLNAQNYMVEIGGEIRAKGKNIEGKPWQIAIEKPTTTGERAVEAVIGLNNMGMASSGDY 254

Query: 262 RIFHVDKNGKRLSHIINPNNKRPISHNLASISVVADSAMTADGLSTGLFVLGETEALKLA 321
            RI+  ++NGKR +H I+P    PI H+LASI+V+A ++MTADGLSTGLFVLGE +AL++A
Sbjct: 255 RIY-FEENGKRFAHEIDPKTGYPIQHHLASITVLAPTSMTADGLSTGLFVLGEDKALEVA 313

Query: 322 EREKLAVFLIVRDKGGYRTAMSSEFEKL 349
            E+  LAV+LI+R  G+ T  SS F+KL
Sbjct: 314 EKNNLAVYLIIRTDNGFVTKSSSAFKKL 341
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 403>:

a111-1.seq

```
     1  ATGCCGTCTG AAACACGCCT GCCGAACTTT ATCCGCACCT TGATATTTGC
    51  CCTGAGTTTT ATCTTCCTGA ACGCCTGTTC GGAACAAACC GCGCAAACCG
   101  TTACCCTGCA AGGTGAAACG ATGGGCACGA CCTATACCGT CAAATACCTT
   151  TCAAATAATC GGGACAAACT CCCCTCACCT GCCGAAATAC AAAAGCGCAT
   201  CGATGACGCG CTTAAAGAAG TCAACCGGCA GATGTCCACC TATCAGCCCG
   251  ACTCCGAAAT CAGCCGGTTC AACCAACACA CAGCCGGCAA GCCCCTCCGC
   301  ATTTCAAGCG ACTTCGCACA CGTTACTGCC GAAGCCGTCC ACCTGAACCG
```

```
 351  CCTGACACAC GGCGCGCTGG ACGTAACCGT CGGCCCCTTG GTCAACCTTT
 401  GGGGATTCGG CCCCGACAAA TCCGTTACCC GTGAACCGTC GCCGGAACAA
 451  ATCAAACAAG CAGCATCTTA TACGGGCATA GACAAAATCA TTTTGAAACA
 501  AGGCAAAGAT TACGCTTCCT TGAGCAAAAC CCACCCCAAG GCCTATTTGG
 551  ATTTATCTTC GATTGCCAAA GGCTTCGGCG TTGATAAAGT TGCGGGCGAA
 601  CTGGAAAAAT ACGGCATTCA AAATTATCTG GTCGAAATCG GCGGCGAGTT
 651  GCACGGCAAA GGCAAAAACG CGCGCGGCGA ACCTTGGCGC ATCGGCATCG
 701  AACAGCCCAA CATCGTCCAA GGCGGCAATA CGCAGATTAT CGTCCCGCTG
 751  AACAACCGTT CGCTTGCCAC TTCCGGCGAT TACCGTATTT TCCACGTCGA
 801  TAAAAGCGGC AAACGCCTCT CCCATATCAT TAATCCGAAC AACAAACGAC
 851  CCATCAGCCA CAACCTCGCC TCCATCAGCG TGGTCGCAGA CAGTGCGATG
 901  ACGGCGGACG GCTTGTCCAC AGGATTATTC GTATTGGGCG AAACCGAAGC
 951  CTTAAAGCTG GCAGAGCGCG AAAAACTCGC TGTTTTCCTG ATTGTCAGGG
1001  ATAAAGGCGG CTACCGCACC GCCATGTCTT CCGAATTTGA AAAACTGCTC
1051  CGCTAA
```

This corresponds to the amino acid sequence <SEQ ID 404; ORF 111-1.a>:

a111-1.pep

```
  1  MPSETRLPNF IRTLIFALSF IFLNACSEQT AQTVTLQGET MGTTYTVKYL
 51  SNNRDKLPSP AEIQKRIDDA LKEVNRQMST YQPDSEISRF NQHTAGKPLR
101  ISSDFAHVTA EAVHLNRLTH GALDVTVGPL VNLWGFGPDK SVTREPSPEQ
151  IKQAASYTGI DKIILKQGKD YASLSKTHPK AYLDLSSIAK GFGVDKVAGE
201  LEKYGIQNYL VEIGGELHGK GKNARGEPWR IGIEQPNIVQ GGNTQIIVPL
251  NNRSLATSGD YRIFHVDKSG KRLSHIINPN NKRPISHNLA SISVVADSAM
301  TADGLSTGLF VLGETEALKL AEREKLAVFL IVRDKGGYRT AMSSEFEKLL
351  R*
```

a111-1/m111-1   98.9% identity in 351 aa overlap

```
                     10        20        30        40        50        60
a111-1.pep   MPSETRLPNFIRTLIFALSFIFLNACSEQTAQTVTLQGETMGTTYTVKYLSNNRDKLPSP
             ||||||||||||:|||||:|||||||||||||||||||||||||||||||||||||||||
m111-1       MPSETRLPNFIRVLIFALGFIFLNACSEQTAQTVTLQGETMGTTYTVKYLSNNRDKLPSP
                     10        20        30        40        50        60

                     70        80        90       100       110       120
a111-1.pep   AEIQKRIDDALKEVNRQMSTYQPDSEISRFNQHTAGKPLRISSDFAHVTAEAVHLNRLTH
             |||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
m111-1       AEIQKRIDDALKEVNRQMSTYQPDSEISRFNQHTAGKPLRISSDFAHVTAEAVRLNRLTH
                     70        80        90       100       110       120

                    130       140       150       160       170       180
a111-1.pep   GALDVTVGPLVNLWGFGPDKSVTREPSPEQIKQAASYTGIDKIILKQGKDYASLSKTHPK
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m111-1       GALDVTVGPLVNLWGFGPDKSVTREPSPEQIKQAASYTGIDKIILKQGKDYASLSKTHPK
                    130       140       150       160       170       180

                    190       200       210       220       230       240
a111-1.pep   AYLDLSSIAKGFGVDKVAGELEKYGIQNYLVEIGGELHGKGKNARGEPWRIGIEQPNIVQ
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m111-1       AYLDLSSIAKGFGVDKVAGELEKYGIQNYLVEIGGELHGKGKNARGEPWRIGIEQPNIVQ
                    190       200       210       220       230       240

                    250       260       270       280       290       300
a111-1.pep   GGNTQIIVPLNNRSLATSGDYRIFHVDKSGKRLSHIINPNNKRPISHNLASISVVADSAM
             |||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
m111-1       GGNTQIIVPLNNRSLATSGDYRIFHVDKNGKRLSHIINPNNKRPISHNLASISVVADSAM
                    250       260       270       280       290       300

                    310       320       330       340       350
a111-1.pep   TADGLSTGLFVLGETEALKLAEREKLAVFLIVRDKGGYRTAMSSEFEKLLRX
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
m111-1       TADGLSTGLFVLGETEALKLAEREKLAVFLIVRDKGGYRTAMSSEFEKLLRX
                    310       320       330       340       350
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 405>:

g114.seq

```
  1  ATGGCTTCCA TCACTTCGCC GCTGCACGGG GCGCAGCAGG AATGCAGCAA
 51  GACTTTTTTA TGTCCGCCGG GCGGGACGAG TATGGGGCGG TCAATGTCGG
```

```
101  TAACGGTAGG TTTGTTTTGT GTTTCCATTA ACTTAACAAT ATCTGTCGAA
151  TACGGTCAAA GCGGCTATTT TACCAGAGCC GCCGAATGTA AAACAGGGTG
201  TCAGGGCATC AGCCCGAGCT GCCTGAACGA ACGGACGGTT TGCGAGGTAA
251  CGATAAAATG GTCGAGCAGC GAAACATCAA CCAGCGACAT GGCCTGTGCC
301  AGCCGCCTTG TGAACATGAT GTCTTCCTGC GAAGGTTCAG GCGAGCCGCC
351  CGGATGGTTG TGCGCGATAA TCAGGCTGTC GGCATATTCG TCCAATGCCA
401  GTTTGACGAT TTCGCGGATG TAA
```

This corresponds to the amino acid sequence <SEQ ID 406; ORF 114.ng>:

```
g114.pep
  1  MASITSPLHG AQQECSKTFL CPPGGTSMGR SMSVTVGLFC VSINLTISVE
 51  YGQSGYFTRA AECKTGCQGI SPSCLNERTV CEVTIKWSSS ETSTSDMACA
101  SRLVNMMSSC EGSGEPPGWL CAIIRLSAYS SNASLTISRM *
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 407>:

```
m114.seq
  1  ATGGCTTCCA TCACTTCGCC GCTGCACGGG GCGCACAGAG AATGCAGCAA
 51  GACTTTTTTA TGTCCACCGG GCGGGACGAG TATAGGGCGG TCAATGTCGG
101  TAACGGTAGG TTTGTTTTGT GTTTCCATTA ACTTAACAAT ATCTGTTGAA
151  TACGGTTGAA GCGGCTATTT TATCAGAGCC GCCGCATGTA AAACAGAGTG
201  TCAGGGCATC AACCCGAGCT GTCTGAACGA ACAGACGCTT TGCGAkGTAA
251  CGATAAAATG GTCGAGCAGC GACACATCGA CCAGCGACAT TGCCTGTGCC
301  AGCCGCCTTG TGAACATGAT GTCTTCCTGC GAArGTTCsG GCGAGCCGcC
351  CGgATGGTTG TGCGCAATAA TCAGGCTGTC GGCATATTCG TCCAATGCCA
401  GTTTGACGAT TTCGCGGATG TAA
```

This corresponds to the amino acid sequence <SEQ ID 408; ORF 114>:

```
m114.pep
  1  MASITSPLHG AHRECSKTFL CPPGGTSIGR SMSVTVGLFC VSINLTISVE
 51  YGXSGYFIRA AACKTECQGI NPSCLNEQTL CXVTIKWSSS DTSTSDIACA
101  SRLVNMMSSC EXSGEPPGWL CAIIRLSAYS SNASLTISRM *
```

```
m114/g114  90.0% identity over a 140 aa overlap
                   10        20        30        40        50        60
m114.pep    MASITSPLHGAHRECSKTFLCPPGGTSIGRSMSVTVGLFCVSINLTISVEYGXSGYFIRA
            |||||||||||::||||||||||||||:|||||||||||||||||||||||| |||| ||
g114        MASITSPLHGAQQECSKTFLCPPGGTSMGRSMSVTVGLFCVSINLTISVEYGQSGYFTRA
                   10        20        30        40        50        60

                   70        80        90       100       110       120
m114.pep    AACKTECQGINPSCLNEQTLCXVTIKWSSSDTSTSDIACASRLVNMMSSCEXSGEPPGWL
            | ||| ||||:||||||:|:| ||||||||:|||||:|||||||||||||| ||||||||
g114        AECKTGCQGISPSCLNERTVCEVTIKWSSSETSTSDMACASRLVNMMSSCEGSGEPPGWL
                   70        80        90       100       110       120

                  130       140
m114.pep    CAIIRLSAYSSNASLTISRMX
            ||||||||||||||||||||
g114        CAIIRLSAYSSNASLTISRMX
                  130       140
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 409>:

a114.seq

```
  1  ATGCCGGAGG CAAGCATCGC CTCCATCACT TCGCCGCTGC ACGGGGCGCA
 51  ACAGGAATGC AGCAAGACTT TTTTATGTCC GCCGGGCGGG ACGAGTATGG
101  GGCGGTCAAT GTCGGTAACG GTAGGTTTGT TTTGTGTTTC CATTAACTTA
151  ACGATATCTG TCGAATACGG TTGAAGCGGC TATTTTATCA GAGCCGCCGC
201  ATGTAAAACA GGGTGTCAGG GCATCAGCCC GAGCTGCCTG AACGAACGGA
251  CGGTTTGCGC CGTTACGATA AAATGGTCGA GCAGCGACAC ATCGACCAGC
301  GACATTGCCT GTGCCAGCCG CCTTGTGAAC ATGATGTCTT CCTGCGAAGG
351  TTCGGGCGAG CCGCCCGGAT GGTTGTGCGC GATAATCAGG CTGTCGGCAT
401  ATTCGTCCAA TGCCAGTTTG ACAATTTCAC GGATGTAA
```

This corresponds to the amino acid sequence <SEQ ID 410; ORF 114.a>:

a114.pep

```
  1  MPEASIASIT SPLHGAQQEC SKTFLCPPGG TSMGRSMSVT VGLFCVSINL
 51  TISVEYG*SG YFIRAAACKT GCQGISPSCL NERTVCAVTI KWSSSDTSTS
101  DIACASRLVN MMSSCEGSGE PPGWLCAIIR LSAYSSNASL TISRM*
```

m114/a114 92.9% identity in 140 aa overlap

```
                            10        20        30        40        50
m114.pep              MASITSPLHGAHRECSKTFLCPPGGTSIGRSMSVTVGLFCVSINLTISVEYGXSG
                      :||||'||||::|||||||||||||.|:||||||||||||||||||||||||||||
a114             MPEASIASITSP!HGAQQECSKTFLCPPGGTSMGRSMSVTVGLFCVSINLTISVEYGXSG
                         10        20        30        40        50        60

              60        70        80        90       100       110
m114.pep     YFIRAAACKTECQGINPSCLNEQTLCXVTIKWSSSDTSTSDIACASRLVNMMSSCEXSGE
             |||||||||| ||||:||||||:|:| ||||||||||||||||||||||||||||| |||
a114         YFIRAAACKTGCQGISPSCLNERTVCAVTIKWSSSDTSTSDIACASRLVNMMSSCEGSGE
                       70        80        90       100       110       120

             120       130       140
m114.pep     PPGWLCAIIRLSAYSSNASLTISRMX
             ||||||||||||||||||||||||||
a114         PPGWLCAIIRLSAYSSNASLTISRMX
                     130       140
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 411>:

g117.seq

```
   1  atggtcgacg aactcgacCT GCTGCCCGAT GCCGTCGCCG CCACCCTGCT
  51  TGCCGACATC GGACGCTACG TCCCCGATTG GAACCTATTG GTTTCCGAGC
 101  GCTGCAACAG CACCGTCGCC GAGCTGGTCA AAGGTGtgga CGAAGTGCAG
 151  AAACTTACCC ACTTCGCCCG GGTGGACAGC CTCGCCACGC CGGAAGAACG
 201  CGCACAGCAA GCGGAAACCA TGCGGAAAAT GCTGCTGGCg atggttaccg
 251  Acatccgcgt cgtaTTAATC AAACTGGCGA TGCGTacgcg caccCTGcta
 301  ttTTtaaGCA ACGCCCCCGA CAGCCCTGAA AAACgcgccG TCgccaaAga
 351  aacccTCGAC ATCTTCGCCC CGCTCGCCAA CCGCTTGGGC GTGTGGCAGC
 401  TCAAATGGCA GCTCGAAGAT TTGGGCTTCC GCCATCAAGA ACCCGAAAAA
 451  TACCGCGAAA TCGCCCTGCT TTTGGACGAA AAACGCACCG AACGCCTCGA
 501  ATACATCGAA AACTTCCTCG ATATCCTGCG TACGGAACTC AAAAAATACA
 551  ATATCCACTT TGAAGTCGCC GGCCGTCCGA AACACATCTA CTCCATTTAC
 601  AAAAAAATGG TGAAGAAAAA ACTCAGCTTC GACGgccTGT TCGACATCCG
 651  CGCCGTGCGG ATTCTGGTCG ATACCGTCCC CGaGTGTTAC ACCACGCTGG
 701  gcaTCGTCCA CAGCCTCTGG CAGCCCATTC CCGGCGagtt CGAcgactAC
 751  ATCGCCAACC CCAAAGgcaA CGgttATAAA AGtTTGCACA CCGTCATCGT
 801  cggcccGGAa gacaaaggtg tggaaGtgCA AATCCGCACC TTCGAtatGC
 851  accAATTCaa CgaatTcggT gtcgccgCCC ACTGGCGtta caaagaaggc
 901  ggcaaaggcg attccGCCtA cgaacaaAAA ATcgccTggt TGCgccaACT
 951  CTTGGACTGG CGCGAAAATA TGGCGGAAAG CGGCAAGGAA GACCTCGCCG
1001  CCGCCTTCAA AACCGAGCTT TTCAACGACA CGATTTATGT TTTGACCCCG
1051  CACGGCAAAG TCCTCTCTCT GCCAACGGGC GCAACCCCCA TCGACTTCGC
1101  CTACGCCCTG CACAGCAGCA TcggCGACCG CTGCCGGGGC GCGAAAGTCG
1151  AaggGCAGAT TGTGCCGCTG TCCACCCCGC TCGAAAACGG ACAGCGCGTC
1201  GAAATcatta cCGCcaaAGA AGGGCATCCT TCCGTCAACT GGCTTTACGA
1251  AGGctgGGtc aAATCCGGCA AGGCCATCGG caaAATCCGC GCCTAcatCC
1301  GCCAGcaaAa cgCcgaCACC GTGCGCGAAG AAGGCCGTGT CCAACTCGAC
1351  AAGCAGCTTG CCAAACTCAC GC[illegible]AAACCC AACCTGCAAG AGCTTgccga
1401  aaATCTCGGC tacaaAAAGC cagaagacct ctacacCGCc gtcggacaag
1451  gcgaaatttc caaccgcgcc atCcaaaaag cctgcggcac GCTgaacgaa
1501  ccgccccCCG TGCCCGTCAG CGCAACCACC ATCGTCAAAC AGTCCAAAAT
```

```
1551  CAAAAAAGGT GGC..AACCG GCGTGCTCAT CGACGGCGAA GACGGCTTGA
1601  TGACCACGCT TGCCAAATGC TGCAAACCCG CGCCGCCCGA CGATATTGCC
1651  GGCTTCGTTA CCCGCGAGCG CGGCATTTCC GTCCACCGCA AAACCTGCCC
1701  CTCTTTCCGA CACCTTGCCG AACACGCGCC CGAAAAAGTA CTGGACGCAA
1751  GTTGGGCGGC GTTGCAGGAA GGGCAAGTGT TCGCCGTCGA TATCGAAATC
1801  CGCGCCCAAG ACCGCTCCGG GCTTTTGCGC GACGTATCCG ACGCGCTCGC
1851  CCGCCACAAA CTCAACGTTA CCGCCGTGCA AACCCAGTCC CGCGACTTGG
1901  AAGCCAGCAT GAGGTTCACG CTCGAAGTCA AACAAGtCAA CGacCTCCCG
1951  CGCGTCCTCG CCGGCCTCGG CGATGTCAAA GGCGTATTGA GCGTTACCCG
2001  GCTTTAA
```

This corresponds to the amino acid sequence <SEQ ID 412; ORF 117.ng>:

```
g117.pep
    1  MVDELDLLPD AVAATLLADI GRYVPDWNLL VSERCNSTVA ELVKGVDEVQ
   51  KLTHFARVDS LATPEERAQQ AETMRKMLLA MVTDIRVVLI KLAMRTRTLL
  101  FLSNAPDSPE KRAVAKETLD IFAPLANRLG VWQLKWQLED LGFRHQEPEK
  151  YREIALLLDE KRTERLEYIE NFLDILRTEL KKYNIHFEVA GRPKHIYSIY
  201  KKMVKKKLSF DGLFDIRAVR ILVDTVPECY TTLGIVHSLW QPIPGEFDDY
  251  IANPKGNGYK SLHTVIVGPE DKGVEVQIRT FDMHQFNEFG VAAHWRYKEG
  301  GKGDSAYEQK IAWLRQLLDW RENMAESGKE DLAAAFKTEL FNDTIYVLTP
  351  HGKVLSLPTG ATPIDFAYAL HSSIGDRCRG AKVEGQIVPL STPLENGQRV
  401  EIITAKEGHP SVNWLYEGWV KSGKAIGKIR AYIRQQNADT VREEGRVQLD
  451  KQLAKLTPKP NLQELAENLG YKKPEDLYTA VGQGEISNRA IQKACGTLNE
  501  PPPVPVSATT IVKQSKIKKG GKTGVLIDGE DGLMTTLAKC CKPAPPDDIA
  551  GFVTRERGIS VHRKTCPSFR HLAEHAPEKV LDASWAALQE GQVFAVDIEI
  601  RAQDRSGLLR DVSDALARHK LNVTAVQTQS RDLEASMRFT LEVKQVNDLP
  651  RVLAGLGDVK GVLSVTRL*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 413>:

```
m117.seq (partial)
    1  ..GTGAAACTCA AGAAATACAA TGTCCATTTC GAAGTCGCCG GCCGCCCGAA
   51    ACACATCTAC TCCATTTACA AAAAAATGGT GAAGAAAAAA CTCAGCTTCG
  101    ACGGCCTCTT TGACATCCGC GCCGTGCGAA TTCTGGTTGA TACCGTCCCC
  151    GAGTGTTACA CCACGCTGGG TATCGTCCAC AGCCTCTGGC AGCCCATTCC
  201    CGGCGAGTTC GACGACTACA TCGCCAATCC CAAAGGCAAC GGCTATAAAA
  251    GTTTGCACAC CGTCATCGTC GGCCCGGAAG ACAAAGGCGT GGAAGTACAA
  301    ATCCGCACCT TCGATATGCA CCAATTCAAC GAATTCGGTG TCGCCGCCCA
  351    CTGgCGTTAC AAAGAGGGCG GCAAGGGCGA TTCCGCCTAC GAACAGAAAA
  401    TCGCCTGGTT GCGCCAACTC TTGGACTGGC GCGAAAACAT GGCGGAAAGC
  451    GGCAAGGAAG ACCTCGCCGC CGCCTTCAAA ACCGAGCTTT TCAACGACAC
  501    GATTTATGTT TTGACCCCGC ACGGCAAAGT CCTCTCCCTG CCCACGGGCG
  551    CGACCCCCAT CGACTTCGCC TACGCCCTGC ACAGCAGCAT CGGCGACCGT
  601    TGCCGCGGTG CGAAAGTCGA AGGGCAGATT GTGCCGCTGT CCACCCCGCT
  651    CGAAAACGGA CAGCGCGTCG AAATCATTAC CGCCAAAGAA GGGCATCCTT
  701    CCGTCAACTG GCTTTACGAA GGCTGGGTCA AATCCAACAA GGCAATCGGC
  751    AAAATCCGCG CCTACATCCG CCAGCAAAAC GCCGACACCG TGCGCGAAGA
  801    AGGCCGCGTC CAACTCGACA AACAGCTTGC CAAACTCACG CCCAAACCCA
  851    ACCTGCAAGA GCTTGCCGAA AATCTCGGCT ACAAAAAGCC AGAAGACCTC
  901    TACACCGCCG TCGGACAAGG CGAAATTTCC AACCGCGCCA TCCAAAAAGC
  951    CTGCGGCACg CTGAACGAAC CGCCGCCCGT ACCCGTCAGC GAAACCACCA
 1001    TCGTCAAACA GTCCAAAATC AAAAAAGGCG GCAAAAACGG CGTGCTCATC
 1051    GACGGCGAAG ACGGTCTGAT GACCACGCTT GCCAAATGCT GCAAACCCGC
 1101    GCCGCCCGAC GATATTATCG GCTTCGTTAC CCGCGAGCGC GgCATTTCAG
 1151    TGCACCGCAA AwyyTkCyCG TCTTTCCAAC ACCTCGCCGA ACACGCGCCC
 1201    GAwAAAGTGC TGGACGCAAG CTGGGCGGCA TTGCAGGAAG GACAAGTATT
 1251    CGCCGTCGAT ATCGAAATCC GCGCCCAAGA CCGCTCCGGG CTTTTGCGCG
 1301    ACGTATCCGA CGCGCTCGCC CGCCACAAAC TCAACGTTAC CGCCGTGCAA
 1351    ACCCAGTCCC GCGACTTGGA AGCCAGCATG AGGTTCACGC TCGAAGTCAA
 1401    ACAAGTCAAC GACCTCCCGC GCGTCCTCGC CAGCCTCGGC GACGTCAAAG
 1451    GCGTATTGAG CGTTACCCGG CTTTAA
```

This corresponds to the amino acid sequence <SEQ ID 414; ORF 117>:

```
m117.pep   (partial)
```

```
  1....VKLKKYNVHF EVAGRPKHIY SIYKKMVKKK LSFDGLFDIR AVRILVDTVP
 51    ECYTTLGIVH SLWQPIPGEF DDYIANPKGN GYKSLHTVIV GPEDKGVEVQ
101    IRTFDMHQFN EFGVAAHWRY KEGGKGDSAY EQKIAWLRQL LDWRENMAES
151    GKEDLAAAFK TELFNDTIYV LTPHGKVLSL PTGATPIDFA YALHSSIGDR
201    CRGAKVEGQI VPLSTPLENG QRVEIITAKE GHPSVNWLYE GWVKSNKAIG
251    KIRAYIRQQN ADTVREEGRV QLDKQLAKLT PKPNLQELAE NLGYKKPEDL
301    YTAVGQGEIS NRAIQKACGT LNEPPPVPVS ETTIVKQSKI KKGGKNGVLI
351    DGEDGLMTTL AKCCKPAPPD DIIGFVTRER GISVHRKXXX SFQHLAEHAP
401    XKVLDASWAA LQEGQVFAVD IEIRAQDRSG LLRDVSDALA RHKLNVTAVQ
451    TQSRDLEASM RFTLEVKQVN DLPRVLASLG DVKGVLSVTR L*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 117 shows 97.6% identity over a 490 aa overlap with a predicted ORF (ORF 117.ng) from *N. gonorrhoeae*:

```
m117/g117

                                                     10        20        30
m117.pep                                  VKLKKYNVHFEVAGRPKHIYSIYKKMVKKKL
                                          :||||:|||||||||||||||||||||||||
g117         EKYREIALLLDEKRTERLEYIENFLDILRTELKKYNIHFEVAGRPKHIYSIYKKMVKKKL
             150       160       170       180       190       200

                     40        50        60        70        80        90
m117.pep     SFDGLFDIRAVRILVDTVPECYTTLGIVHSLWQPIPGEFDDYIANPKGNGYKSLHTVIVG
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g117         SFDGLFDIRAVRILVDTVPECYTTLGIVHSLWQPIPGEFDDYIANPKGNGYKSLHTVIVG
             210       220       230       240       250       260

                    100       110       120       130       140       150
m117.pep     PEDKGVEVQIRTFDMHQFNEFGVAAHWRYKEGGKGDSAYEQKIAWLRQLLDWRENMAESG
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g117         PEDKGVEVQIRTFDMHQFNEFGVAAHWRYKEGGKGDSAYEQKIAWLRQLLDWRENMAESG
             270       280       290       300       310       320

                    160       170       180       190       200       210
m117.pep     KEDLAAAFKTELFNDTIYVLTPHGKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGQIV
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g117         KEDLAAAFKTELFNDTIYVLTPHGKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGQIV
             330       340       350       360       370       380

                    220       230       240       250       260       270
m117.pep     PLSTPLENGQRVEIITAKEGHPSVNWLYEGWVKSNKAIGKIRAYIRQQNADTVREEGRVQ
             |||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
g117         PLSTPLENGQRVEIITAKEGHPSVNWLYEGWVKSGKAIGKIRAYIRQQNADTVREEGRVQ
             390       400       410       420       430       440

                    280       290       300       310       320       330
m117.pep     LDKQLAKLTPKPNLQELAENLGYKKPEDLYTAVGQGEISNRAIQKACGTLNEPPPVPVSE
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
g117         LDKQLAKLTPKPNLQELAENLGYKKPEDLYTAVGQGEISNRAIQKACGTLNEPPPVPVSA
             450       460       470       480       490       500

                    340       350       360       370       380       390
m117.pep     TTIVKQSKIKKGGKNGVLIDGEDGLMTTLAKCCKPAPPDDIIGFVTRERGISVHRKXXXS
             ||||||||||||||:||||||||||||||||||||||||||| ||||||||||||||:  |
g117         TTIVKQSKIKKGGKTGVLIDGEDGLMTTLAKCCKPAPPDDIAGFVTRERGISVHRKTCPS
             510       520       530       540       550       560

                    400       410       420       430       440       450
m117.pep     FQHLAEHAPXKVLDASWAALQEGQVFAVDIEIRAQDRSGLLRDVSDALARHKLNVTAVQT
```

```
             |:||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g117         FRHLAEHAPEKVLDASWAALQEGQVFAVDIEIRAQDRSGLLRDVSDALARHKLNVTAVQT
            570       580       590       600       610       620

                     460       470       480       490
m117.pep     QSRDLEASMRFTLEVKQVNDLPRVLASLGDVKGVLSVTRLX
             |||||||||||||||||||||||||||:||||||||||||||
g117         QSRDLEASMRFTLEVKQVNDLPRVLAGLGDVKGVLSVTRLX
            630       640       650       660
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 415>:

a117.seq

```
   1  ATGGTTCATG AACTCGACCT GCTCCCCGAT GCCGTCGCCG CCACCCTGCT
  51  TGCCGACATC GGACGCTACG TCCCCGACTG GAACCTATTG GTTTCCGAAC
 101  GCTGCAACAG TACCGTCGCC GAGCTGGTCA AAGGTGTGGA CGAAGTGCAG
 151  AAACTCACCC ACTTCGCCCG GGTGGACAGC CTCGCCACGC CGGAAGAACG
 201  CGCCCAGCAG GCAGAAACTA TGCGGAAAAT GCTGCTGGCG ATGGTTACCG
 251  ACATCCGCGT CGTGTTAATC AAACTGGCGA TGCGTACGCG CACCCTGCAA
 301  TTTTTAAGCA ACGCCCCCGA CAGCCCCGAA AAACGCGCCG TCGCCAAAGA
 351  AACCCTCGAC ATCTTCGCCC CGCTCGCCAA CCGTTTGGGC GTGTGGCAGC
 401  TCAAATGGCA GCTCGAAGAT TTGGGCTTCC GCCATCAAGA ACCCGAAAAA
 451  TACCGCGAAA TCGCCCTGCT TTTGGACGAA AAACGCACCG AACGCCTCGA
 501  ATACATCGAA AACTTCCTTA ATATCCTGCG TACGGAACTC AAAAAATACA
 551  ATATCCACTT TGAAGTCGCC GGCCGTCCGA AACACATCTA CTCCATTTAC
 601  AAAAAAATGG TGAAGAAAAA ACTCAGCTTC GACGGGTTGT TCGACATCCG
 651  CGCCGTGCGG ATTCTGGTTG ATACCGTCCC CGAGTGTTAC ACCACACTGG
 701  GCATTGTCCA CAGCCTCTGG CAGCCCATTC CCGGCGAGTT CGACGACTAC
 751  ATCGCCAACC CGAAAGGCAA CGGCTATAAA AGTTTGCACA CCGTCATCGT
 801  CGGCCCGGAA GACAAAGGCG TGGAAGTGCA AATCCGCACC TTCGATATGC
 851  ACCAATTCAA CGAATTCGGT GTCGCCGCGC ACTGGCGTTA CAAAGAGGGC
 901  GGCAAAGGCG ATTCCGCCTA CGAACAAAAA ATCGCCTGGT TACGCCAACT
 951  TTTGGACTGG CGCGAAAACA TGGCGGAAAG CGGCAAGGAA GACCTCGCCG
1001  CCGCCTTCAA AACCGAGCTT TTCAACGACA CGATTTATGT TTTGACCCCG
1051  CACGGCAAAG TCCTCTCCCT GCCCACAGGC GCGACCCCCA TCGACTTCGC
1101  CTACGCCCTG CACAGCAGCA TCGGCGACCG TTGCCGCGGT GCGAAAGTCG
1151  AAGGGCAGAT TGTGCCGCTG TCCACCCCGC TCGAAAACGG ACAGCGTGTC
1201  GAAATCATTA CCGCCAAAGA AGGGCATCCT TCCGTCAACT GGCTTTACGA
1251  AGGCTGGGTC AAATCCAACA AGGCAATCGG CAAAATCCGC GCCTACATCC
1301  GCCAGCAAAA CGCCGACACC GTGCGCGAAG AAGGCCGCGT CCAACTCGAC
1351  AAACAGCTTG CCAAACTCAC GCCCAAACCC AACCTGCAAG AGCTTGCCGA
1401  AAATCTCGGC TACAAAAAGC CAGAAGACCT CTACACCGCC GTCGGACAAG
1451  GCGAAATTTC CAACCGCGCC ATCCAAAAAG CCTGCGGCAC GCTGAACGAA
1501  CCGCCGCCCG TACCCGTCAG CGAAACCACC ATCGTCAAAC AGTCCAAAAT
1551  CAAAAAAGGC GGCAAAAACG GCGTGCTCAT CGACGGCGAA GACGGTCTGA
1601  TGACCACGCT TGCCAAATGC TGCAAACCCG CGCCGCCCGA CGACATTGTC
1651  GGCTTCGTTA CCCGCGATCG CGGCATTTCG GTACACCGCA AAACCTGCCC
1701  CTCTTTCCGA CACCTCGCCG AACACGCGCC CGAAAAAGTA CTGGACGCAA
1751  GTTGGGCGGC GTTGCAGGAA GGACAAGTGT TCGCCGTCGA TATCGAAATC
1801  CGCGCCCAAG ACCGCTCCGG GCTTTTGCGC GACGTATCCG ACGCGCTCGC
1851  CCGCCACAAA CTCAACGTTA CCGCCGTGCA AACCCAGTCC CGCGACTTGG
1901  AAGCCAGCAT GAGGTTCACG CTCGAAGTCA AACAAGTTAC CGACCTCCCA
1951  CGCGTCCTCG CCAGCCTCGG CGACGTCAAA GGCGTATTGA GCGTTACCCG
2001  GCTTTAA
```

This corresponds to the amino acid sequence <SEQ ID 416; ORF 117.a>:

a117.pep

```
  1  MVHELDLLPD AVAATLLADI GRYVPDWNLL VSERCNSTVA ELVKGVDEVQ
 51  KLTHFARVDS LATPEERAQQ AETMRKMLLA MVTDIRVVLI KLAMRTRTLQ
101  FLSNAPDSPE KRAVAKETLD IFAPLANRLG VWQLKWQLED LGFRHQEPEK
151  YREIALLLDE KRTERLEYIE NFLNILRTEL KKYNIHFEVA GRPKHIYSIY
201  KKMVKKKLSF DGLFDIRAVR ILVDTVPECY TTLGIVHSLW QPIPGEFDDY
251  IANPKGNGYK SLHTVIVGPE DKGVEVQIRT FDMHQFNEFG VAAHWRYKEG
301  GKGDSAYEQK IAWLRQLLDW RENMAESGKE DLAAAFKTEL FNDTIYVLTP
351  HGKVLSLPTG ATPIDFAYAL HSSIGDRCRG AKVEGQIVPL STPLENGQRV
```

```
401  EIITAKEGHP SVNWLYEGWV KSNKAIGKIR AYIRQQNADT VREEGRVQLD
451  KQLAKLTPKP NLQELAENLG YKKPEDLYTA VGQGEISNRA IQKACGTLNE
501  PPPVPVSETT IVKQSKIKKG GKNGVLIDGE DGLMTTLAKC CKPAPPDDIV
551  GFVTRDRGIS VHRKTCPSFR HLAEHAPEKV LDASWAALQE GQVFAVDIEI
601  RAQDRSGLLR DVSDALARHK LNVTAVQTQS RDLEASMRFT LEVKQVTDLP
651  RVLASLGDVK GVLSVTRL*
```

m117/a117 98.0% identity in 490 aa overlap

```
                                            10        20        30
m117.pep                             VKLKKYNVHFEVAGRPKHIYSIYKKMVKKKL
                                     :.||||:||||||||||||||||||||||||
a117        EKYREIALLLDEKRTERLEYIENFLNILRTELKKYNIHFEVAGRPKHIYSIYKKMVKKKL
           150       160       170       180       190       200

                    40        50        60        70        80        90
m117.pep    SFDGLFDIRAVRILVDTVPECYTTLGIVHSLWQPIPGEFDDYIANPKGNGYKSLHTVIVG
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a117        SFDGLFDIRAVRILVDTVPECYTTLGIVHSLWQPIPGEFDDYIANPKGNGYKSLHTVIVG
           210       220       230       240       250       260

                   100       110       120       130       140       150
m117.pep    PEDKGVEVQIRTFDMHQFNEFGVAAHWRYKEGGKGDSAYEQKIAWLRQLLDWRENMAESG
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a117        PEDKGVEVQIRTFDMHQFNEFGVAAHWRYKEGGKGDSAYEQKIAWLRQLLDWRENMAESG
           270       280       290       300       310       320

                   160       170       180       190       200       210
m117.pep    KEDLAAAFKTELFNDTIYVLTPHGKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGQIV
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a117        KEDLAAAFKTELFNDTIYVLTPHGKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGQIV
           330       340       350       360       370       380

                   220       230       240       250       260       270
m117.pep    PLSTPLENGQRVEIITAKEGHPSVNWLYEGWVKSNKAIGKIRAYIRQQNADTVREEGRVQ
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a117        PLSTPLENGQRVEIITAKEGHPSVNWLYEGWVKSNKAIGKIRAYIRQQNADTVREEGRVQ
           390       400       410       420       430       440

                   280       290       300       310       320       330
m117.pep    LDKQLAKLTPKPNLQELAENLGYKKPEDLYTAVGQGEISNRAIQKACGTLNEPPPVPVSE
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a117        LDKQLAKLTPKPNLQELAENLGYKKPEDLYTAVGQGEISNRAIQKACGTLNEPPPVPVSE
           450       460       470       480       490       500

                   340       350       360       370       380       390
m117.pep    TTIVKQSKIKKGGKNGVLIDGEDGLMTTLAKCCKPAPPDDIIGFVTRERGISVHRKXXXS
            ||||||||||||||||||||||||||||||||||||||||:|||||:||||||||:   |
a117        TTIVKQSKIKKGGKNGVLIDGEDGLMTTLAKCCKPAPPDDIVGFVTRDRGISVHRKTCPS
           510       520       530       540       550       560

                   400       410       420       430       440       450
m117.pep    FQHLAEHAPXKVLDASWAALQEGQVFAVDIEIRAQDRSGLLRDVSDALARHKLNVTAVQT
            |:||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
a117        FRHLAEHAPEKVLDASWAALQEGQVFAVDIEIRAQDRSGLLRDVSDALARHKLNVTAVQT
           570       580       590       600       610       620

                   460       470       480       490
m117.pep    QSRDLEASMRFTLEVKQVNDLPRVLASLGDVKGVLSVTRLX
            ||||||||||||||||||:|''|||||||||||||||||||
a117        QSRDLEASMRFTLEVKQVTDLPRVLASLGDVKGVLSVTRLX
           630       640       650       660
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 417>:

g117-1.seq

```
   1 ATGACCGCCA TCAGCCCGAT TCAAGACACG CAAAGCGCGA CCCTGCAAGA
  51 ATTGCGCGAA TGGTTCGACA GCTACTGCGC CGCTCTGCCG GACAACGATA
 101 AAAACCTCAT CGGTACCGCA TGGTCGCTGG CGCAGGAACA TTATCCTGCC
 151 GATGCCGCCA CGCCGTATGG CGAGCCGCTG CCCGACCACT TCCTCGGCGC
 201 GGCGCAAATG GTCGACGAAC TCGACCTGCT GCCCGATGCC GTCGCCGCCA
 251 CCCTGCTTGC CGACATCGGA CGCTACGTCC CCGATTGGAA CCTATTGGTT
 301 TCCGAGCGCT GCAACAGCAC CGTCGCCGAG CTGGTCAAAG GTGTGGACGA
 351 AGTGCAGAAA CTTACCCACT TCGCCCGGGT GGACAGCCTC GCCACGCCGG
 401 AAGAACGCGC ACAGCAAGCG GAAACCATGC GGAAAATGCT GCTGGCGATG
 451 GTTACCGACA TCCGCGTCGT ATTAATCAAA CTGGCGATGC GTACGCGCAC
 501 CCTGCAATTT TTAAGCAACG CCCCCGACAG CCCTGAAAAA CGCGCCGTCG
 551 CCAAAGAAAC CCTCGACATC TTCGCCCCGC TCGCCAACCG CTTGGGCGTG
 601 TGGCAGCTCA AATGGCAGCT CGAAGATTTG GGCTTCCGCC ATCAAGAACC
 651 CGAAAAATAC CGCGAAATCG CCCTGCTTTT GGACGAAAAA CGCACCGAAC
 701 GCCTCGAATA CATCGAAAAC TTCCTCGATA TCCTGCGTAC GGAACTCAAA
 751 AAATACAATA TCCACTTTGA AGTCGCCGGC CGTCCGAAAC ACATCTACTC
 801 CATTTACAAA AAAATGGTGA AGAAAAAACT CAGCTTCGAC GGCCTGTTCG
 851 ACATCCGCGC CGTGCGGATT CTGGTCGATA CCGTCCCCGA GTGTTACACC
 901 ACGCTGGGCA TCGTCCACAG CCTCTGGCAG CCCATTCCCG GCGagttCGA
 951 cgactACATC GCCAACCCCA AAGgcaACGg ttATAAAAGt TTGCACACCG
1001 TCATCGTcgg cccGGAagaa aaaggtgtgg aagtgcAAAT CCGCACCTTC
1051 GATATGCacc AATTCaaCga ATTCGGTGTC GCCGCCCACT GGCGTTACAA
1101 AGAAGGCGGC AAAGGCGATT CCGCCTACGA ACAAAAAATC GCCTGGTTGC
1151 GCCAACTCTT GGACTGGCGC GAAAATATGG CGGAAAGCGG CAAGGAAGAC
1201 CTCGCCGCCG CCTTCAAAAC CGAGCTTTTC AACGACACGA TTTATGTTTT
1251 GACCCCGCAC GGCAAAGTCC TCTCTCTGCC AACGGGCGCA ACCCCCATCG
1301 ACTTCGCCTA CGCCCTGCAC AGCAGCATCG GCGACCGCTG CCGGGGCGCG
1351 AAAGTCGAAG GGCAGATTGT GCCGCTGTCC ACCCCGCTCG AAAACGGACA
1401 GCGCGTCGAA ATCATTACCG CCAAAGAAGG GCATCCTTCC GTCAACTGGC
1451 TTTACGAAGG CTGGGTCAAA TCCGGCAAGG CCATCGGCAA AATCCGCGCC
1501 TACATCCGCC AGCAAAACGC CGACACCGTG CGCGAAGAAG GCCGTGTCCA
1551 ACTCGACAAG CAGCTTGCCA AACTCACGCC CAAACCCAAC CTGCAAGAGC
1601 TTGCCGAAAA TCTCGGCTAC AAAAAGCCAG AAGACCTCTA CACCGCCGTC
1651 GGACAAGGCG AAATTTCCAA CCGCGCCATC CAAAAAGCCT GCGGCACGCT
1701 GAACGAACCG CCGCCCGTGC CCGTCAGCGC AACCACCATC GTCAAACAGT
1751 CCAAAATCAA AAAAGGTGGC AAAACCGGCG TGCTCATCGA CGGCGAAGAC
1801 GGCTTGATGA CCACGCTTGC CAAATGCTGC AAACCCGCGC CGCCCGACGA
1851 TATTGCCGGC TTCGTTACCC GCGAGCGCGG CATTTCCGTC CACCGCAAAA
1901 CCTGCCCCTC TTTCCGACAC CTTGCCGAAC ACGCGCCCGA AAAAGTACTG
1951 GACGCAAGTT GGGCGGCGTT GCAGGAAGGG CAAGTGTTCG CCGTCGATAT
2001 CGAAATCCGC GCCCAAGACC GCTCCGGGCT TTTGCGCGAC GTATCCGACG
2051 CGCTCGCCCG CCACAAACTC AACGTTACCG CCGTGCAAAC CCAGTCCCGC
2101 GACTTGGAAG CCAGCATGAG GTTCACGCTC GAAGTCAAAC AAGTCAACGA
2151 CCTCCCGCGC GTCCTCGCCG GCCTCGGCGA TGTCAAAGGC GTATTGAGCG
2201 TTACCCGGCT TTAA
```

This corresponds to the amino acid sequence <SEQ ID 418; ORF 117-1.ng>:

g117-1.pep

```
   1 MTAISPIQDT QSATLQELRE WFDSYCAALP DNDKNLIGTA WSLAQEHYPA
  51 DAATPYGEPL PDHFLGAAQM VDELDLLPDA VAATLLADIG RYVPDWNLLV
 101 SERCNSTVAE LVKGVDEVQK LTHFARVDSL ATPEERAQQA ETMRKMLLAM
 151 VTDIRVVLIK LAMRTRTLQF LSNAPDSPEK RAVAKETLDI FAPLANRLGV
 201 WQLKWQLEDL GFRHQEPEKY REIALLLDEK RTERLEYIEN FLDILRTELK
 251 KYNIHFEVAG RPKHIYSIYK KMVKKKLSFD GLFDIRAVRI LVDTVPECYT
 301 TLGIVHSLWQ PIPGEFDDYI ANPKGNGYKS LHTVIVGPEE KGVEVQIRTF
 351 DMHQFNEFGV AAHWRYKEGG KGDSAYEQKI AWLRQLLDWR ENMAESGKED
 401 LAAAFKTELF NDTIYVLTPH GKVLSLPTGA TPIDFAYALH SSIGDRCRGA
 451 KVEGQIVPLS TPLENGQRVE IITAKEGHPS VNWLYEGWVK SGKAIGKIRA
 501 YIRQQNADTV REEGRVQLDK QLAKLTPKPN LQELAENLGY KKPEDLYTAV
 551 GQGEISNRAI QKACGTLNEP PPVPVSATTI VKQSKIKKGG KTGVLIDGED
 601 GLMTTLAKCC KPAPPDDIAG FVTRERGISV HRKTCPSFRH LAEHAPEKVL
 651 DASWAALQEG QVFAVDIEIR AQDRSGLLRD VSDALARHKL NVTAVQTQSR
 701 DLEASMRFTL EVKQVNDLPR VLAGLGDVKG VLSVTRL*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 419>:

m117-1.seq

```
   1 ATGACCGCCA TCAGCCCGAT TCAAGACACG CAAAGCGCGA CTCTGCAAGA
```

```
  51  ATTGCGCGAA TGGTTCGACA GCTACTGCGC CGCTCTGCCG GACAACGATA
 101  AAAACCTCAT CGGTACCGCA TGGTTGCTGG CGCAGGAACA TTACCCCGCC
 151  GATGCCGCCA CGCCGTATGG CGAGCCGCTG CCCGACCACT TCCTCGGCGC
 201  GGCGCAAATG GTTCATGAAC TCGACCTGCT CCCCGATGCC GTCGCCGCCA
 251  CCCTGCTTGC CGACATCGGA CGCTACGTCC CCGACTGGAA CCTATTGGTT
 301  TCCGAACGCT GCAACAGTAC CGTCGCCGAG CTGGTCAAAG GTGTGGACGA
 351  AGTGCAGAAA CTCACCCACT TCGCCCGGGT GGACAGCCTC GCCACGCCGG
 401  AAGAACGCGC CCAGCAGGCA GAAACTATGC GGAAAATGCT GCTGGCGATG
 451  GTTACCGACA TCCGCGTCGT GTTAATCAAA CTGGCGATGC GTACGCGCAC
 501  CCTGCAATTT TTAAGCAACG CCCCCGACAG CCCCGAAAAA CGCGCCGTCG
 551  CCAAAGAAAC CCTCGACATC TTCGCCCCGC TCGCCAACCG TTTGGGCGTG
 601  TGGCAGCTCA AATGGCAGCT CGAAGATTTG GGCTTCCGCC ATCAAAAGCC
 651  CGAAAAATAC CGCGAAATCG CGCTGCTTTT GGACGAAAAA CTCACCGAAC
 701  GCCTCGAATA CATCGAAAAC TTCCTCAACA TCCTGCGCGG TGAACTCAAG
 751  AAATACAATG TCCATTTCGA AGTCGCCGGC CGCCCGAAAC ACATCTACTC
 801  CATTTACAAA AAAATGGTGA AGAAAAAACT CAGCTTCGAC GGCCTCTTTG
 851  ACATCCGCGC CGTGCGAATT CTGGTTGATA CCGTCCCCGA GTGTTACACC
 901  ACGCTGGGTA TCGTCCACAG CCTCTGGCAG CCCATTCCCG GCGAGTTCGA
 951  CGACTACATC GCCAATCCCA AAGGCAACGG CTATAAAAGT TTGCACACCG
1001  TCATCGTCGG CCCGGAAGAC AAAGGCGTGG AAGTACAAAT CCGCACCTTC
1051  GATATGCACC AATTCAACGA ATTCGGTGTC GCCGCCCACT GGCGTTACAA
1101  AGAGGGCGGC AAGGGCGATT CCGCCTACGA ACAGAAAATC GCCTGGTTGC
1151  GCCAACTCTT GGACTGGCGC GAAAACATGG CGGAAAGCGG CAAGGAAGAC
1201  CTCGCCGCCG CCTTCAAAAC CGAGCTTTTC AACGACACGA TTTATGTTTT
1251  GACCCCGCAC GGCAAAGTCC TCTCCCTGCC CACGGGCGCG ACCCCCATCG
1301  ACTTCGCCTA CGCCCTGCAC AGCAGCATCG GCGACCGTTG CCGCGGTGCG
1351  AAAGTCGAAG GGCAGATTGT GCCGCTGTCC ACCCCGCTCG AAAACGGACA
1401  GCGCGTCGAA ATCATTACCG CCAAAGAAGG GCATCCTTCC GTCAACTGGC
1451  TTTACGAAGG CTGGGTCAAA TCCAACAAGG CAATCGGCAA AATCCGCGCC
1501  TACATCCGCC AGCAAAACGC CGACACCGTG CGCGAAGAAG GCCGCGTCCA
1551  ACTCGACAAA CAGCTTGCCA AACTCACGCC CAAACCCAAC CTGCAAGAGC
1601  TTGCCGAAAA TCTCGGCTAC AAAAAGCCAG AAGACCTCTA CACCGCCGTC
1651  GGACAAGGCG AAATTTCCAA CCGCGCCATC CAAAAAGCCT GCGGCACGCT
1701  GAACGAACCG CCGCCCGTAC CCGTCAGCGA AACCACCATC GTCAAACAGT
1751  CCAAAATCAA AAAAGGCGGC AAAAACGGCG TGCTCATCGA CGGCGAAGAC
1801  GGTCTGATGA CCACGCTTGC CAAATGCTGC AAACCCGCGC CGCCCGACGA
1851  TATTATCGGC TTCGTTACCC GCGAGCGCGG CATTTCAGTG CACCGCAAAA
1901  CCTGCCCGTC TTTCCAACAC CTCGCCGAAC ACGCGCCCGA AAAAGTGCTG
1951  GACGCAAGCT GGGCGGCATT GCAGGAAGGA CAAGTATTCG CCGTCGATAT
2001  CGAAATCCGC GCCCAAGACC GCTCCGGGCT TTTGCGCGAC GTATCCGACG
2051  CGCTCGCCCG CCACAAACTC AACGTTACCG CCGTGCAAAC CCAGTCCCGC
2101  GACTTGGAAG CCAGCATGAG GTTCACGCTC GAAGTCAAAC AAGTCAACGA
2151  CCTCCCGCGC GTCCTCGCCA GCCTCGGCGA CGTCAAAGGC GTATTGAGCG
2201  TTACCCGGCT TTAA
```

This corresponds to the amino acid sequence <SEQ ID 420; ORF 117-1>:

m117-1.pep

```
  1  MTAISPIQDT QSATLQELRE WFDSYCAALP DNDKNLIGTA WLLAQEHYPA
 51  DAATPYGEPL PDHFLGAAQM VHELDLLPDA VAATLLADIG RYVPDWNLLV
101  SERCNSTVAE LVKGVDEVQK LTHFARVDSL ATPEERAQQA ETMRKMLLAM
151  VTDIRVVLIK LAMRTRTLQF LSNAPDSPEK RAVAKETLDI FAPLANRLGV
201  WQLKWQLEDL GFRHQKPEKY REIALLLDEK RTERLEYIEN FLNILRGELK
251  KYNVHFEVAG RPKHIYSIYK KMVKKKLSFD GLFDIRAVRI LVDTVPECYT
301  TLGIVHSLWQ PIPGEFDDYI ANPKGNGYKS LHTVIVGPED KGVEVQIRTF
351  DMHQFNEFGV AAHWRYKEGG KGDSAYEQKI AWLRQLLDWR ENMAESGKED
401  LAAAFKTELF NDTIYVLTPH GKVLSLPTGA TPIDFAYALH SSIGDRCRGA
451  KVEGQIVPLS TPLENGQRVE IITAKEGHPS VNWLYEGWVK SNKAIGKIRA
501  YIRQQNADTV REEGRVQLDK QLAKLTPKPN LQELAENLGY KKPEDLYTAV
551  GQGEISNRAI QKACGTLNEP PPVPVSETTI VKQSKIKKGG KNGVLIDGED
601  GLMTTLAKCC KPAPPDDIIG FVTRERGISV HRKTCPSFQH LAEHAPEKVL
651  DASWAALQEG QVFAVDIEIR AQDRSGLLRD VSDALARHKL NVTAVQTQSR
701  DLEASMRFTL EVKQVNDLPR VLASLGDVKG VLSVTRL*
```

m117-1/g117-1   98.2% identity in 737 aa overlap

```
                    10        20        30        40        50        60
m117-1.pep  MTAISPIQDTQSATLQELREWFDSYCAALPDNDKNLIGTAWLLAQEHYPADAATPYGEPL
            |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
g117-1      MTAISPIQDTQSATLQELREWFDSYCAALPDNDKNLIGTAWSLAQEHYPADAATPYGEPL
                    10        20        30        40        50        60

                    70        80        90       100       110       120
```

```
m117-1.pep  PDHFLGAAQMVHELDLLPDAVAATLLADIGRYVPDWNLLVSERCNSTVAELVKGVDEVQK
            |||||||||| ||||||||||||||||||||||||||||||'|||||||||||||||||
g117-1      PDHFLGAAQMVDELDLLPDAVAATLLADIGRYVPDWNLLVSERCNSTVAELVKGVDEVQK
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m117-1.pep  LTHFARVDSLATPEERAQQAETMRKMLLAMVTDIRVVLIKLAMRTRTLQFLSNAPDSPEK
            ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
g117-1      LTHFARVDSLATPEERAQQAETMRKMLLAMVTDIRVVLIKLAMRTRTLQFLSNAPDSPEK
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m117-1.pep  RAVAKETLDIFAPLANRLGVWQLKWQLEDLGFRHQKPEKYREIALLLDEKRTERLEYIEN
            |||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
g117-1      RAVAKETLDIFAPLANRLGVWQLKWQLEDLGFRHQEPEKYREIALLLDEKRTERLEYIEN
                   190       200       210       220       230       240

                   250       260       270       280       290       300
m117-1.pep  FLNILRGELKKYNVHFEVAGRPKHIYSIYKKMVKKKLSFDGLFDIRAVRILVDTVPECYT
            ||:||| |||||:|||||||||||||||||||||||||||||||||||||||||||||||
g117-1      FLDILRTELKKYNIHFEVAGRPKHIYSIYKKMVKKKLSFDGLFDIRAVRILVDTVPECYT
                   250       260       270       280       290       300

                   310       320       330       340       350       360
m117-1.pep  TLGIVHSLWQPIPGEFDDYIANPKGNGYKSLHTVIVGPEDKGVEVQIRTFDMHQFNEFGV
            ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
g117-1      TLGIVHSLWQPIPGEFDDYIANPKGNGYKSLHTVIVGPEEKGVEVQIRTFDMHQFNEFGV
                   310       320       330       340       350       360

                   370       380       390       400       410       420
m117-1.pep  AAHWRYKEGGKGDSAYEQKIAWLRQLLDWRENMAESGKEDLAAAFKTELFNDTIYVLTPH
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g117-1      AAHWRYKEGGKGDSAYEQKIAWLRQLLDWRENMAESGKEDLAAAFKTELFNDTIYVLTPH
                   370       380       390       400       410       420

                   430       440       450       460       470       480
m117-1.pep  GKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGQIVPLSTPLENGQRVEIITAKEGHPS
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g117-1      GKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGQIVPLSTPLENGQRVEIITAKEGHPS
                   430       440       450       460       470       480

                   490       500       510       520       530       540
m117-1.pep  VNWLYEGWVKSNKAIGKIRAYIRQQNADTVREEGRVQLDKQLAKLTPKPNLQELAENLGY
            ||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
g117-1      VNWLYEGWVKSGKAIGKIRAYIRQQNADTVREEGRVQLDKQLAKLTPKPNLQELAENLGY
                   490       500       510       520       530       540

                   550       560       570       580       590       600
m117-1.pep  KKPEDLYTAVGQGEISNRAIQKACGTLNEPPPVPVSETTIVKQSKIKKGGKNGVLIDGED
            |||||||||||||||||||||||||||||||||||| |||||||||||||||:|||||||
g117-1      KKPEDLYTAVGQGEISNRAIQKACGTLNEPPPVPVSATTIVKQSKIKKGGKTGVLIDGED
                   550       560       570       580       590       600

                   610       620       630       640       650       660
m117-1.pep  GLMTTLAKCCKPAPPDDIIGFVTRERGISVHRKTCPSFQHLAEHAPEKVLDASWAALQEG
            |||||||||||||||||| |||||||||||||||||||:|||||||||||||||||||||
g117-1      GLMTTLAKCCKPAPPDDIAGFVTRERGISVHRKTCPSFRHLAEHAPEKVLDASWAALQEG
                   610       620       630       640       650       660

                   670       680       690       700       710       720
m117-1.pep  QVFAVDIEIRAQDRSGLLRDVSDALARHKLNVTAVQTQSRDLEASMRFTLEVKQVNDLPR
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g117-1      QVFAVDIEIRAQDRSGLLRDVSDALARHKLNVTAVQTQSRDLEASMRFTLEVKQVNDLPR
                   670       680       690       700       710       720

                   730
m117-1.pep  VLASLGDVKGVLSVTRLX
            |||:||||||||||||||
g117-1      VLAGLGDVKGVLSVTRLX
                   730
```

m117-1/RelA

```
sp|P55133|RELA_VIBSS GTP PYROPHOSPHOKINASE (ATP:GTP 3'-PYROPHOSPHOTRANSFERASE) (PPGPP
SYNTHETASE I) >gi|537617 (U13769) ppGpp synthetase I (Vibrio sp.) Length = 744
 Score =  536 bits (1366), Expect = e-151
 Identities = 288/685 (42%), Positives = 432/685 (63%), Gaps = 31/685 (4%)

Query: 74  LDLLPDAVAATLLADI---GRYVPDWNLLVSERCNSTVAELVKGVDEVQKLTHFARVDSL 130
            L + D + A LL +   G Y D    + E + T+ LV+GV+++ -+   ++ S
Sbjct: 68  LSMDADTLIAALLYPLVEGGCYSTD---ALKEEYSGTILHLVQGVEQMCAIS---QLKST 121

Query: 131 ATPEERAQQAETMRKMLLAMVTDIRVVLIKLAMRTRTLQFLSNAPDSPEKRAVAKETLDI 190
            A    +A Q + +R+MLL+MV D R V+IKLA R  L+ + + PD  +RA A+E +I
Sbjct: 122 AEETAQAAQVDNIRRMLLSMVDDFRCVVIKLAERICNLREVKDCPDEV-RRAAAQECANI 180

Query: 191 FAPLANRLGVWQLKWQLEDLGFRHQKPEKYREIALLLDEKRTERLEYIENFLNILRGELK 250
            +APLANRLG+ QLKW++ED  FR+Q P+ Y++IA  L E+R +R +YI +F++ L   +K
Sbjct: 181 YAPLANRLGIGQLKWEIEDYAFRYQHPDTYKQIAKQLSERRIDREDYITHFVDDLSDAMK 240

Query: 251 KYNVHFEVAGRPKHIYSIYKKMVKKKLSFDGLFDIRAVRILVDTVPECYTTLGIVHSLWQ 310
             N+  EV GRPKHIYSI++KM KK L FD LFD+RAVRI+ + + +CY  LG+VH+ ++
Sbjct: 241 ASNIRAEVQGRPKHIYSIWRKMQKKSLEFDELFDVRAVRIVAEELQDCYAALGVVHTKYR 300

Query: 311 PIPGEFDDYIANPKGNGYKSLHTVIVGPEDKGVEVQIRTFDMHQFNEFGVAAHWRYKEG- 369
            +P EFDDY+ANPK NGY+S+HTV++GPE K +E+QIRT  MH+ +E GVAAHW+YKEG
Sbjct: 301 HLPKEFDDYVANPKPNGYQSIHTVVLGPEGKTIEIQIRTKQMHEESELGVAAHWKYKEGT 360

Query: 370 --GKGDSAYEQKIAWLRQLLDWRENMAESGKEDLAAAFKTELFNDTIYVLTPHGKVLSLP 427
              G  SAY++KI WLR+LL W+E M++SG  ++    ++++F+D +Y  TP G V+ LP
Sbjct: 361 ASGGAQSAYDEKINWLRKLLAWQEEMSDSG--EMLDELRSQVFDDRVYAFTPKGDVVDLP 418

Query: 428 TGATPIDFAYALHSSIGDRCRGAKVEGQIVPLSTPLENGQRVEIITAKEGHPSVNWLYE- 486
            + ATP+DFAY +HS +G RC GAKVEG+IVP +  L+ G +VEIIT KE +PS +WL
Sbjct: 419 SNATPLDFAYHIHSEVGHRCIGAKVEGRIVPFTYHLQMGDQVEIITQKEPNPSRDWLNPN 478

Query: 487 -GWVKSNKAIGKIRAYIRQQNADTVREEGRVQLDKQLAKL--TPKPNLQELAENLGYKKP 543
            G+V S++A  K+ A+ R+Q+ D     G+  L+ +L K+  T K      A+    K P
Sbjct: 479 LGFVTSSRARAKVHAWFRKQDRDKNIIAGKEILEAELVKIHATLKDAQYYAAKRFNVKSP 538

Query: 544 EDLYTAVGQGEIS-NRAIQKACGTLNEPPPVPVSETTIVKQSKI--------KKGGKNGV 594
            E+LY  +G G++  N+ I     +N+P        +  + K S+         KK  +- V
Sbjct: 539 EELYAGIGSGDLRINQVINHINALVNKPTAEEEDQQLLEKLSEASNKQATSHKKPQRDAV 598

Query: 595 LIDGEDGLMTTLAKCCKPAPPDDIIGFVTRERGISVHRKTCPSFQHLAEHAPEKVLDASW 654
            +++G D LMT LA+CC+P P DDI GFVT+ RGISVHR  C   + L  HAPE++D W
Sbjct: 599 VVEGVDNLMTHLARCCQPIPGDDIQGFVTQGRGISVHRMDCEQLEELRHHAPERIIDTVW 658

Query: 655 AALQEGQVFAVDIEIRAQDRSGLLRDVSDALARHKLNVTAVQTQ--SRDLEASMRFTLEV 712
              G  + + + + A +R+GLL++++ L   K+ V  ++++  +     + M F LE+
Sbjct: 659 GGGFVGN-YTITVRVTASERNGLLKELTNTLMNEKVKVAGMKSRVDYKKQMSIMDFELEL 717

Query: 713 KQVNDLPRVLASLGDVKGVLSVTRL 737
              + L RVL  +  VK V    RL
Sbjct: 718 TDLEVLGRVLKRIEQVKDVAEAKRL 742
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 421>:

a117-1.seq

```
  1  ATGACCGCCA TCAGCCCGAT TCAAGACACG CAAAGCGCGA CTCTGCAAGA
 51  ATTGCGCGAA TGGTTCGACA GCTACTGCAC CGCGCTGCCG AACAACGATA
101  AAAAACTTGT CTTAGCCGCC CGTTCGCTGG CGGAAGCACA TTACCCCGCC
151  GATGCCGCCA CGCCGTATGG CGAACCGCTG CCCGACCACT TCCTCGGCGC
201  GGCGCAAATG GTTCATGAAC TCGACCTGCT CCCCGATGCC GTCGCCGCCA
251  CCCTGCTTGC CGACATCGGA CGCTACGTCC CCGACTGGAA CCTATTGGTT
301  TCCGAACGCT GCAACAGTAC CGTCGCCGAG CTGGTCAAAG GTGTGGACGA
351  AGTGCAGAAA CTCACCCACT TCGCCCGGGT GGACAGCCTC GCCACGCCGG
401  AAGAACGCGC CCAGCAGGCA GAAACTATGC GGAAAATGCT GCTGGCGATG
451  GTTACCGACA TCCGCGTCGT GTTAATCAAA CTGGCGATGC GTACGCGCAC
501  CCTGCAATTT TTAAGCAACG CCCCCGACAG CCCCGAAAAA CGCGCCGTCG
551  CCAAAGAAAC CCTCGACATC TTCGCCCCGC TCGCCAACCG TTTGGGCGTG
601  TGGCAGCTCA AATGGCAGCT CGAAGATTTG GG.TTCCGCC ATCAAGAACC
651  CGAAAAATAC CGCGAAATCG CCCTGCTTTT GGACGAAAAA CGCACCGAAC
701  GCCTCGAATA CATCGAAAAC TTCCTTAATA TCCTGCGTAC GGAACTCAAA
751  AAATACAATA TCCACTTTGA AGTCGCCGGC CGTCCGAAAC ACATCTACTC
801  CATTTACAAA AAAATGGTGA AGAAAAAACT CAGCTTCGAC GGGTTGTTCG
```

```
 851  ACATCCGCGC CGTGCGGATT CTGGTTGATA CCGTCCCCGA GTGTTACACC
 901  ACACTGGGCA TTGTCCACAG CCTCTGGCAG CCCATTCCCG GCGAGTTCGA
 951  CGACTACATC GCCAACCCGA AAGGCAACGG CTATAAAAGT TTGCACACCG
1001  TCATCGTCGG CCCGGAAGAC AAAGGCGTGG AAGTGCAAAT CCGCACCTTC
1051  GATATGCACC AATTCAACGA ATTCGGTGTC GCCGCGCACT GGCGTTACAA
1101  AGAGGGCGGC AAAGGCGATT CCGCCTACGA ACAAAAAATC GCCTGGTTAC
1151  GCCAACTTTT GGACTGGCGC GAAAACATGG CGGAAAGCGG CAAGGAAGAC
1201  CTCGCCGCCG CCTTCAAAAC CGAGCTTTTC AACGACACGA TTTATGTTTT
1251  GACCCCGCAC GGCAAAGTCC TCTCCCTGCC CACAGGCGCG ACCCCCATCG
1301  ACTTCGCCTA CGCCCTGCAC AGCAGCATCG GCGACCGTTG CCGCGGTGCG
1351  AAAGTCGAAG GGCAGATTGT GCCGCTGTCC ACCCCGCTCG AAAACGGACA
1401  GCGTGTCGAA ATCATTACCG CCAAAGAAGG GCATCCTTCC GTCAACTGGC
1451  TTTACGAAGG CTGGGTCAAA TCCAACAAGG CAATCGGCAA AATCCGCGCC
1501  TACATCCGCC AGCAAAACGC CGACACCGTG CGCGAAGAAG GCCCCGTCCA
1551  ACTCGACAAA CAGCTTGCCA AACTCACGCC CAAACCCAAC CTGCAAGAGC
1601  TTGCCGAAAA TCTCGGCTAC AAAAAGCCAG AAGACCTCTA CACCGCCGTC
1651  GGACAAGGCG AAATTTCCAA CCGCGCCATC CAAAAAGCCT GCGGCACGCT
1701  GAACGAACCG CCGCCCGTAC CCGTCAGCGA AACCACCATC GTCAAACAGT
1751  CCAAAATCAA AAAAGGCGGC AAAAACGGCG TGCTCATCGA CGGCGAAGAC
1801  GGTCTGATGA CCACGCTTGC CAAATGCTGC AAACCCGCGC CGCCCGACGA
1851  CATTGTCGGC TTCGTTACCC GCGATCGCGG CATTTCGGTA CACCGCAAAA
1901  CCTGCCCCTC TTTCCGACAC CTCGCCGAAC ACGCGCCCGA AAAAGTACTG
1951  GACGCAAGTT GGGCGGCGTT GCAGGAAGGA CAAGTGTTCG CCGTCGATAT
2001  CGAAATCCGC GCCCAAGACC GCTCCGGGCT TTTGCGCGAC GTATCCGACG
2051  CGCTCGCCCG CCACAAACTC AACGTTACCG CCGTGCAAAC CCAGTCCCGC
2101  GACTTGGAAG CCAGCATGAG GTTCACGCTC GAAGTCAAAC AAGTTACCGA
2151  CCTCCCACGC GTCCTCGCCA GCCTCGGCGA CGTCAAAGGC GTATTGAGCG
2201  TTACCCGGCT TTAA
```

This corresponds to the amino acid sequence <SEQ ID 422; ORF 117-1.a>:

a117-1.pep

```
  1  MTAISPIQDT QSATLQELRE WFDSYCTALP NNDKKLVLAA RSLAEAHYPA
 51  DAATPYGEPL PDHFLGAAQM VHELDLLPDA VAATLLADIG RYVPDWNLLV
101  SERCNSTVAE LVKGVDEVQK LTHFARVDSL ATPEERAQQA ETMRKMLLAM
151  VTDIRVVLIK LAMRTRTLQF LSNAPDSPEK RAVAKETLDI FAPLANRLGV
201  WQLKWQLEDL GFRHQEPEKY REIALLLDEK RTERLEYIEN FLNILRTELK
251  KYNIHFEVAG RPKHIYSIYK KMVKKKLSFD GLFDIRAVRI LVDTVPECYT
301  TLGIVHSLWQ PIPGEFDDYI ANPKGNGYKS LHTVIVGPED KGVEVQIRTF
351  DMHQFNEFGV AAHWRYKEGG KGDSAYEQKI AWLRQLLDWR ENMAESGKED
401  LAAAFKTELF NDTIYVLTPH GKVLSLPTGA TPIDFAYALH SSIGDRCRGA
451  KVEGQIVPLS TPLENGQRVE IITAKEGHPS VNWLYEGWVK SNKAIGKIRA
501  YIRQQNADTV REEGRVQLDK QLAKLTPKPN LQELAENLGY KKPEDLYTAV
551  GQGEISNRAI QKACGTLNEP PPVPVSETTI VKQSKIKKGG KNGVLIDGED
601  GLMTTLAKCC KPAPPDDIVG FVTRDRGISV HRKTCPSFRH LAEHAPEKVL
651  DASWAALQEG QVFAVDIEIR AQDRSGLLRD VSDALARHKL NVTAVQTQSR
701  DLEASMRFTL EVKQVTDLPR VLASLGDVKG VLSVTRL*
```

a117-1/m117-1 97.7% identity in 737 aa overlap

```
                     10        20        30        40        50        60
m117-1.pep   MTAISPIQDTQSATLQELREWFDSYCAALPDNDKNLIGTAWLLAQEHYPADAATPYGEPL
             ||||||||||||||||||||||||||:|||:|||:|: :|  ||: ||||||||||||||
a117-1       MTAISPIQDTQSATLQELREWFDSYCTALPNNDKKLVLAARSLAEAHYPADAATPYGEPL
                     10        20        30        40        50        60

                     70        80        90       100       110       120
m117-1.pep   PDHFLGAAQMVHELDLLPDAVAATLLADIGRYVPDWNLLVSERCNSTVAELVKGVDEVQK
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a117-1       PDHFLGAAQMVHELDLLPDAVAATLLADIGRYVPDWNLLVSERCNSTVAELVKGVDEVQK
                     70        80        90       100       110       120

                    130       140       150       160       170       180
m117-1.pep   LTHFARVDSLATPEERAQQAETMRKMLLAMVTDIRVVLIKLAMRTRTLQFLSNAPDSPEK
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a117-1       LTHFARVDSLATPEERAQQAETMRKMLLAMVTDIRVVLIKLAMRTRTLQFLSNAPDSPEK
                    130       140       150       160       170       180

                    190       200       210       220       230       240
m117-1.pep   RAVAKETLDIFAPLANRLGVWQLKWQLEDLGFRHQKPEKYREIALLLDEKRTERLEYIEN
             |||||||||||||||||||||||||||||||||||:||||||||||.||||||||||||
a117-1       RAVAKETLDIFAPLANRLGVWQLKWQLEDLGFRHQEPEKYREIALLLDEKRTERLEYIEN
                    190       200       210       220       230       240
```

```
                250       260       270       280       290       300
m117-1.pep  FLNILRGELKKYNVHFEVAGRPKHIYSIYKKMVKKKLSFDGLFDIRAVRILVDTVPECYT
            |||||| ||||||:|||||||||||||||||||||||||||||||||||||||||||||||
a117-1      FLNILRTELKKYNIHFEVAGRPKHIYSIYKKMVKKKLSFDGLFDIRAVRILVDTVPECYT
                250       260       270       280       290       300

                310       320       330       340       350       360
m117-1.pep  TLGIVHSLWQPIPGEFDDYIANPKGNGYKSLHTVIVGPEDKGVEVQIRTFDMHQFNEFGV
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a117-1      TLGIVHSLWQPIPGEFDDYIANPKGNGYKSLHTVIVGPEDKGVEVQIRTFDMHQFNEFGV
                310       320       330       340       350       360

                370       380       390       400       410       420
m117-1.pep  AAHWRYKEGGKGDSAYEQKIAWLRQLLDWRENMAESGKEDLAAAFKTELFNDTIYVLTPH
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a117-1      AAHWRYKEGGKGDSAYEQKIAWLRQLLDWRENMAESGKEDLAAAFKTELFNDTIYVLTPH
                370       380       390       400       410       420

                430       440       450       460       470       480
m117-1.pep  GKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGQIVPLSTPLENGQRVEIITAKEGHPS
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a117-1      GKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGQIVPLSTPLENGQRVEIITAKEGHPS
                430       440       450       460       470       480

                490       500       510       520       530       540
m117-1.pep  VNWLYEGWVKSNKAIGKIRAYIRQQNADTVREEGRVQLDKQLAKLTPKPNLQELAENLGY
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a117-1      VNWLYEGWVKSNKAIGKIRAYIRQQNADTVREEGRVQLDKQLAKLTPKPNLQELAENLGY
                490       500       510       520       530       540

                550       560       570       580       590       600
m117-1.pep  KKPEDLYTAVGQGEISNRAIQKACGTLNEPPPVPVSETTIVKQSKIKKGGKNGVLIDGED
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a117-1      KKPEDLYTAVGQGEISNRAIQKACGTLNEPPPVPVSETTIVKQSKIKKGGKNGVLIDGED
                550       560       570       580       590       600

                610       620       630       640       650       660
m117-1.pep  GLMTTLAKCCKPAPPDDIIGFVTRERGISVHRKTCPSFQHLAEHAPEKVLDASWAALQEG
            ||||||||||||||||||:|||||:|||||||||||||:|||||||||||||||||||||
a117-1      GLMTTLAKCCKPAPPDDIVGFVTRDRGISVHRKTCPSFRHLAEHAPEKVLDASWAALQEG
                610       620       630       640       650       660

                670       680       690       700       710       720
m117-1.pep  QVFAVDIEIRAQDRSGLLRDVSDALARHKLNVTAVQTQSRDLEASMRFTLEVKQVNDLPR
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
a117-1      QVFAVDIEIRAQDRSGLLRDVSDALARHKLNVTAVQTQSRDLEASMRFTLEVKQVTDLPR
                670       680       690       700       710       720

                730
m117-1.pep  VLASLGDVKGVLSVTRLX
            ||||||||||||||||||
a117-1      VLASLGDVKGVLSVTRLX
                730
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 423>:

```
g118.seq
      1  ATGTGCGAGT TCAAGGATTT TAGAAGAAAC ATCCCTTGTT TTGAAGAGTA
     51  TGACGAAAAT TCATTTATTG GCAAATGGTA TGATGACGGG GTGTGGGATG
    101  ATGAAGAATA TTGGAAGCTG GAGAATGATT TAATcgaGGT TAGGAGAAAA
    151  TATCCTTATC CGATGGATAT ACCAAGGGAT ATTGTGATTG GAATCGGTAC
    201  CATTATTGAT TTTTTAATGG TTCCAAATTG GGAGCTTTTT GAAATTAAAG
    251  CTTCCCCTTG GTTGCCTGAT AGCGTGGGAA TTCATGAACG TTATGAAAGA
    301  TTCACAACGA TGCTCCGTTA TATTTTTACC GAGAAAGACA TAGTCAACGT
    351  GCGATTTGAT TATTACAaCA AAAAATAG
```

This corresponds to the amino acid sequence <SEQ ID 424; ORF 118.ng>:

```
g118.pep
      1  MCEFKDFRRN IPCFEEYDEN SFIGKWYDDG VWDDEEYWKL ENDLIEVRRK
     51  YPYPMDIPRD IVIGIGTIID FLMVPNWELF EIKASPWLPD SVGIHERYER
    101  FTTMLRYIFT EKDIVNVRFD YYNKK*
```

```
                           70         80         90        100        110        120
m118.pep          VVIGIGTIIDFLMVPNWKLFEIKASPWLPDSVGIHERYERFTTMLRYIFTEKDIVNVRFD
                  :||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
a118              IVIGIGTIIDFLMVPNWELFEIKASPWLPDSVGIHERYERFTTMLRYIFTEKDIVNVRFD
                           70         80         90        100        110        120

m118.pep          YYNKKX
                  ||||||
a118              YYNKKX
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 429>:

```
g120.seq
      1  ATGATGAAGA CTTTTAAAAA TATATTTTCC GCCGCCATTT TGTCCGCCGC
     51  CCTGCCGTGC GCGTATGCGG CAAGGCTACC CCAATCCGCC GTGCTGCACT
    101  ATTCCGGCAG CTACGGCATT CCCGCCACGA TGACATTTGA ACGCAGCGGC
    151  AATGCTTACA AAATCGTTTC GACGATTAAA GTGCCGCTAT ACAATATCCG
    201  TTTCGAATCC GGCGGTACGG TTGTCGGCAA TACCCTGCAC CCTGCCTACT
    251  ATAAAGACAT ACGCAGGGGC AAACTGTATG CGGAAGCCAA ATTCGCCGAC
    301  GGCAGCGTAA CCTACGGCAA AGCGGGCGAG AGCAAAACCG AGCAAAGCCC
    351  CAAGGCTATG GATTTGTTCA CGCTTGCCTG GCAGTTGGCG GCAAATGACG
    401  CGAAACTCCC CCCGGGTCTG AAAATCACCA ACGGCAAAAA ACTTTATTCC
    451  GTCGGCGGCC TGAATAAGGC GGGTACGGGA AAATACAGCA Taggcggcgt
    501  gGAAACCGAA GTCGTCAAAT ATCGGGTGCG GCGCGGCGAC GATACGGTAA
    551  CGTATTTCTT CGCACCGTCC CTGAACAATA TTCCGGCACA AATCGGCTAT
    601  ACCGAcgaCG GCAAAACCTA TACGCTGAAG CTCAAATCGG TGCAGATCAA
    651  CGGACAGGCC GCCAAACCGT AA
```

This corresponds to the amino acid sequence <SEQ ID 430; ORF 120.ng>:

```
g120.pep
      1  MMKTFKNIFS AAILSAALPC AYAARLPQSA VLHYSGSYGI PATMTFERSG
     51  NAYKIVSTIK VPLYNIRFES GGTVVGNTLH PAYYKDIRRG KLYAEAKFAD
    101  GSVTYGKAGE SKTEQSPKAM DLFTLAWQLA ANDAKLPPGL KITNGKKLYS
    151  VGGLNKAGTG KYSIGGVETE VVKYRVRRGD DTVTYFFAPS LNNIPAQIGY
    201  TDDGKTYTLK LKSVQINGQA AKP*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 431>:

```
m120.seq
      1  ATGATGAAGA CTTTTAAAAA TATATTTTCC GCCGCCATTT TGTCCGCCGC
     51  CCTGCCGTGC GCGTATGCGG CAGGGCTGCC CCAATCCGCC GTGCTGmACT
    101  ATTCCGGCAG CTACGGCATT CCCGCCACGA TGACATTTGA ACGCAGCGGC
    151  AATGCTTACA AAATCGTTTC GACGATTAAA GTGCCGCTAT ACAATATCCG
    201  TTTCGAGTCC GGCGGTACGG TTGTCGGCAA TACCCTGCAC CCTACCTACT
    251  ATAGAGACAT ACGCAGGGGC AAACTGTATG CGGAAGCcAA ATTCGCCGAC
    301  GGCAGCGTAA CTTACGGCAA AGCGGGCGAG AGCAAAACCG AGCAAAGCCC
    351  CAAGGCTATG GATTTGTTCA CGCTTGCCTG GCAGTTGGCG GCAAATGACG
    401  CGAAACTCCC CCCGGGGCTG AAAATCACCA ACGGCAAAAA ACTTTATTCC
    451  GTCGGCGGTT TGAATAAGGC GGGTACAGGA AAATACAGCA TAGGCGGCGT
    501  GGAAACCGAA GTCGTCAAAT ATCGGGTGCG GCGCGGCGAC GATGCGGTAA
    551  TGTATTTCTT CGCACCGTCC CTGAACAATA TTCCGGCACA AATCGGCTAT
    601  ACCGACGACG GCAAAACCTA TACGCTGAAA CTCAAATCGG TGCAGATCAA
    651  CGGCCAGGCA GCCAAACCG
```

This corresponds to the amino acid sequence <SEQ ID 432; ORF 120>:

```
m120.pep
      1  MMKTFKNIFS AAILSAALPC AYAAGLPQSA VLXYSGSYGI PATMTFERSG
     51  NAYKIVSTIK VPLYNIRFES GGTVVGNTLH PTYYRDIRRG KLYAEAKFAD
    101  GSVTYGKAGE SKTEQSPKAM DLFTLAWQLA ANDAKLPPGL KITNGKKLYS
    151  VGGLNKAGTG KYSIGGVETE VVKYRVRRGD DAVMYFFAPS LNNIPAQIGY
    201  TDDGKTYTLK LKSVQINGQA AKP
```

```
                  10        20        30        40        50        60

                  70        80        90       100       110       120
m120.pep   VPLYNIRFESGGTVVGNTLHPTYYRDIRRGKLYAEAKFADGSVTYGKAGESKTEQSPKAM
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a120       VPLYNIRFESGGTVVGNTLHPTYYRDIRRGKLYAEAKFADGSVTYGKAGESKTEQSPKAM
                  70        80        90       100       110       120

                 130       140       150       160       170       180
m120.pep   DLFTLAWQLAANDAKLPPGLKITNGKKLYSVGGLNKAGTGKYSIGGVETEVVKYRVRRGD
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a120       DLFTLAWQLAANDAKLPPGLKITNGKKLYSVGGLNKAGTGKYSIGGVETEVVKYRVRRGD
                 130       140       150       160       170       180

                 190       200       210       220
m120.pep   DAVMYFFAPSLNNIPAQIGYTDDGKTYTLKLKSVQINGQAAKPX
           ||||||||||||||||||||||||||||||||||||||||||||
a120       DAVMYFFAPSLNNIPAQIGYTDDGKTYTLKLKSVQINGQAAKPX
                 190       200       210       220
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 435>:

g121.seq

```
   1  ATGGAAACAC AGCTTTACAT CGGCATTATG TCGGGAACCA GTATGGACGG
  51  GGCGGATGCC GTGCTGGTAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
 101  AAGGGCACGC CTTTACCCCC TACCCTGACC GGTTGCGCCG CAAATTGCTG
 151  GATTTGCAGG ACACAGGCAC AGACGAACTG CACCGCAGCA GGATGTTGTC
 201  GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
 251  GTCAAAACCT CGCTCCGTGC GACATTACCG CCCTCGGCTG CCACGGGCAA
 301  ACCGTCCGAC ACGCGCCGGA ACACGGTtac AGCATACAGC TTGCCGATTT
 351  GCCGCTGCTG GCGGAACTGa cgcggatttT TACCGTCggc gacttcCGCA
 401  GCCGCGACCT TGCTGCCGGC GGacaAGGTG CGCCGCTCGT CCCCGCCTTT
 451  CACGAAGCCC TGTTCCGCGA TGACAGGGAA ACACGCGTGG TACTGAACAT
 501  CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCGGCGCA CCCGCCTTCG
 551  GCTTCGACAC AGGGCCGGGC AATATGCTGA TGGAcgcgtg gacgcaggca
 601  cacTGGcagc TGCCTTACGA CAAAAacggt gcAAAGgcgg cacAAGGCAA
 651  catatTGCcg cAACTGCTCG gcaggctGCT CGCCcacccG TATTTCTCAC
 701  AACCCcaccc aaAAAGCACG GGgcGCGaac TgtttgcccT AAattggctc
 751  gaaacctAcc ttgacggcgg cgaaaaccga tacgacgtat tgcggacgct
 801  ttcccgattc accgcgcaaA ccgTttggga cgccgtctca CACGCAGCGG
 851  CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
 901  TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
 951  CACCGCCGAA CTGAACCTCG ATCCTCAATG GGTGGAGGCG gccgCATTTg
1001  cgtggttggC GGCGTGTTGG ATTAACCGCA TTCCCGGTAG TCCGCACAAA
1051  GCGACCGGCG CATCCAAACC GTGTATTCTG GGCGCGGGAT ATTATTATTG
1101  A
```

This corresponds to the amino acid sequence <SEQ ID 436; ORF 121.ng>:

g121.pep

```
   1  METQLYIGIM SGTSMDGADA VLVRMDGGKW LGAEGHAFTP YPDRLRRKLL
  51  DLQDTGTDEL HRSRMLSQEL SRLYAQTAAE LLCSQNLAPC DITALGCHGQ
 101  TVRHAPEHGY SIQLADLPLL AELTRIFTVG DFRSRDLAAG GQGAPLVPAF
 151  HEALFRDDRE TRVVLNIGGI ANISVLPPGA PAFGFDTGPG NMLMDAWTQA
 201  HWQLPYDKNG AKAAQGNILP QLLGRLLAHP YFSQPHPKST GRELFALNWL
 251  ETYLDGGENR YDVLRTLSRF TAQTVWDAVS HAAADARQMY ICGGGIRNPV
 301  LMADLAECFG TRVSLHSTAE LNLDPQWVEA AAFAWLAACW INRIPGSPHK
 351  ATGASKPCIL GAGYYY*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 437>:

m121.seq

```
   1  ATGGAAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
  51  GGCGGATGCC GTACTGATAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
 101  AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTTACGCCG CCAATTGCTG
```

```
                    310       320       330       340       350       360
m121.pep     LMADLAECFGTRVSLHSTADLNLDPQWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL
             ||||||||||||||||||:|||||||||||| ||||||||||||||||||||||||||||
g121         LMADLAECFGTRVSLHSTAELNLDPQWVEAAAFAWLAACWINRIPGSPHKATGASKPCIL
                    310       320       330       340       350       360

m121.pep     XAGYYYX
              ||||||
g121         GAGYYYX
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 439>:

```
a121.seq
       1  ATGGAAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
      51  GGCGGATGCC GTACTGATAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
     101  AAGGGCACGC CTTTACCCCC TACCCCGGC GGTTACGCCG CAAATTGCTG
     151  GATTTGCAGG ACACAGGCGC GGACGAACTG CACCGCAGCA GGATGTTGTC
     201  GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
     251  GTCAAAACCT CGCGCCGTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
     301  ACCGTCAGAC ACGCGCCGGA ACACAGTTAC AGCGTACAGC TTGCCGATTT
     351  GCCGCTGCTG GCGGAACGGA CTCAGATTTT TACCGTCGGC GACTTCCGCA
     401  GCCGCGACCT TGCGGCCGGC GGACAAGGCG CGCCGCTCGT CCCCGCCTTT
     451  CACGAAGCCC TGTTCCGCGA CGACAGGGAA ACACGCGCGG TACTGAACAT
     501  CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCGACGCA CCCGCCTTCG
     551  GCTTCGACAC AGGACCGGGC AATATGCTGA TGGACGCGTG GATGCAGGCA
     601  CACTGGCAGC TTCCTTACGA CAAAAACGGT GCAAAGGCGG CACAAGGCAA
     651  CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCCACCCG TATTTCGCAC
     701  AACCCCACCC TAAAAGCACG GGGCGCGAAC TGTTTGCCCT AAATTGGCTC
     751  GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TGCGGACGCT
     801  TTCCCGATTC ACCGCGCAAA CCGTTTTCGA CGCCGTCTCA CACGCAGCGG
     851  CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
     901  TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
     951  CACCGCCGAA CTGAACCTCG ATCCGCAATG GGTAGAAGCC GCCGCGTTCG
    1001  CATGGATGGC GGCGTGTTGG GTCAACCGCA TTCCCGGTAG TCCGCACAAA
    1051  GCAACCGGCG CATCCAAACC GTGTATTCTG GGCGCGGGAT ATTATTATTG
    1101  A
```

This corresponds to the amino acid sequence <SEQ ID 440; ORF 121.a>:

```
a121.pep
       1  METQLYIGIM SGTSMDGADA VLIRMDGGKW LGAEGHAFTP YPGRLRRKLL
      51  DLQDTGADEL HRSRMLSQEL SRLYAQTAAE LLCSQNLAPS DITALGCHGQ
     101  TVRHAPEHSY SVQLADLPLL AERTQIFTVG DFRSRDLAAG GQGAPLVPAF
     151  HEALFRDDRE TRAVLNIGGI ANISVLPPDA PAFGFDTGPG NMLMDAWMQA
     201  HWQLPYDKNG AKAAQGNILP QLLDRLLAHP YFAQPHPKST GRELFALNWL
     251  ETYLDGGENR YDVLRTLSRF TAQTVFDAVS HAAADARQMY ICGGGIRNPV
     301  LMADLAECFG TRVSLHSTAE LNLDPQWVEA AAFAWMAACW VNRIPGSPHK
     351  ATGASKPCIL GAGYYY*
```

m121/a121 74.0% identity in 366 aa overlap

```
                    10        20        30        40        50        60
m121.pep     METQLYIGIMSGTSMDGADAVLIRMDGGKWLGAEGHAFTPYPGRLRRQLLDLQDTGADEL
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
a121         METQLYIGIMSGTSMDGADAVLIRMDGGKWLGAEGHAFTPYPGRLRRKLLDLQDTGADEL
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m121.pep     HRSRILSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPEHGYSIQLADLPLL
             ||||:|||||||||||||||||||||||||||||||||||||||||||:||:||||||||
a121         HRSRMLSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPEHSYSVQLADLPLL
                    70        80        90       100       110       120
```

```
                      10        20        30        40        50        60
m121-1.pep  METQLYIGIMSGTSMDGADAVLIRMDGGKWLGAEGHAFTPYPGRLRRQLLDLQDTGADEL
            ||||||||||||||||||||||:|||||||||||||||||||| ||||:||||||||:|||
g121        METQLYIGIMSGTSMDGADAVLVRMDGGKWLGAEGHAFTPYPDRLRRKLLDLQDTGTDEL
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m121-1.pep  HRSRILSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPEHGYSIQLADLPLL
            ||||:||||||||||||||||||||||||| |||||||||||||||||||||||||||||
g121        HRSRMLSQELSRLYAQTAAELLCSQNLAPCDITALGCHGQTVRHAPEHGYSIQLADLPLL
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m121-1.pep  AERTRIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDNRETRAVLNIGGIANISVLPPDA
            || ||||||||||||||||||||||||||||||||||:||||:|||||||||||||||| |
g121        AELTRIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDDRETRVVLNIGGIANISVLPPGA
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m121-1.pep  PAFGFDTGPGNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST
            ||||||||||||||||||||||||||||||||||||||||||| ||||||||:|||||||
g121        PAFGFDTGPGNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLGRLLAHPYFSQPHPKST
                   190       200       210       220       230       240

                   250       260       270       280       290       300
m121-1.pep  GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVCDAVSHAAADARQMYICGGGIRNPV
            ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
g121        GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVWDAVSHAAADARQMYICGGGIRNPV
                   250       260       270       280       290       300

                   310       320       330       340       350       360
m121-1.pep  LMADLAECFGTRVSLHSTADLNLDPQWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL
            |||||||||||||||||||:||||||||||| ||||||||||||||||||||||||||||
g121        LMADLAECFGTRVSLHSTAELNLDPQWVEAAAFAWLAACWINRIPGSPHKATGASKPCIL
                   310       320       330       340       350       360

m121-1.pep  XAGYYYX
             ||||||
g121        GAGYYYX
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 443>:

a121-1.seq

```
   1  ATGGAAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
  51  GGCGGATGCC GTACTGATAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
 101  AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTTACGCCG CAAATTGCTG
 151  GATTTGCAGG ACACAGGCGC GGACGAACTG CACCGCAGCA GGATGTTGTC
 201  GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
 251  GTCAAAACCT CGCGCCGTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
 301  ACCGTCAGAC ACGCGCCGGA ACACAGTTAC AGCGTACAGC TTGCCGATTT
 351  GCCGCTGCTG GCGGAACGGA CTCAGATTTT TACCGTCGGC GACTTCCGCA
 401  GCCGCGACCT TGCGGCCGGC GGACAAGGCG CGCCGCTCGT CCCCGCCTTT
 451  CACGAAGCCC TGTTCCGCGA CGACAGGGAA ACACGCGCGG TACTGAACAT
 501  CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCGACGCA CCCGCCTTCG
 551  GCTTCGACAC AGGACCGGGC AATATGCTGA TGGACGCGTG GATGCAGGCA
 601  CACTGGCAGC TTCCTTACGA CAAAAACGGT GCAAAGGCGG CACAAGGCAA
 651  CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCCACCCG TATTTCGCAC
 701  AACCCCACCC TAAAAGCACG GGGCGCGAAC TGTTTGCCCT AAATTGGCTC
 751  GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TGCGGACGCT
 801  TTCCCGATTC ACCGCGCAAA CCGTTTTCGA CGCCGTCTCA CACGCAGCGG
 851  CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
 901  TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
 951  CACCGCCGAA CTGAACCTCG ATCCGCAATG GGTAGAAGCC GCCGCGTTCG
1001  CATGGATGGC GGCGTGTTGG GTCAACCGCA TTCCCGGTAG TCCGCACAAA
1051  GCAACCGGCG CATCCAAACC GTGTATTCTG GGCGCGGGAT ATTATTATTG
1101  A
```

This corresponds to the amino acid sequence <SEQ ID 444; ORF 121-1.a>:

a121-1.pep

```
   1  METQLYIGIM SGTSMDGADA VLIRMDGGKW LGAEGHAFTP YPGRLRRKLL
  51  DLQDTGADEL HRSRMLSQEL SRLYAQTAAE LLCSQNLAPS DITALGCHGQ
 101  TVRHAPEHSY SVQLADLPLL AERTQIFTVG DFRSRDLAAG GQGAPLVPAF
 151  HEALFRDDRE TRAVLNIGGI ANISVLPPDA PAFGFDTGPG NMLMDAWMQA
```

```
g122.pep
      1  MALLSIRKLH KQYGSVTAIQ SLDLDLEKGE VIVLLGPSGC GKSTLLRCVN
     51  GLEPHQGGSI VMDGVGEFGK DVSWQTARQK VGMVFQSNEL FAHMTVIENI
    101  FLGPVKEQNR DRAEAEAQAG KLLERVGLLD RKNAYPRELS GGQKQRIAIV
    151  RALCLNPEVI LLDEITAALD PEMVREVLEV VLELAREGMS MLIVTHEMGF
    201  ARKVADRIVF MDKGGIVESS DPETFFSAPK SERARQFLAG MDY*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 447>:

```
m122.seq
      1  GTTGTCATGA TTAAAATCCG CAATATCCAT AAGACCTTTG GCGAAAACAC
     51  TATTTTGCGC GGCATCGATT TGGATGTGTG CAAAGGGCAG GTGGTCGTCA
    101  TCCTCGGGcC TTCCGGCTCA GGCAAAACGA CGTTTCTGCG ATGCCTAAAC
    151  GCGTTGGAAA TGCCCGAAGA CGGACAAATC GAGTTCGACA ACGAGCGACC
    201  GCTGAAAATC GATTTTTCTA AAAAACCAAG CAAACACGAT ATTTTGGCAC
    251  TGCGCCGCAA ATCAkGCATG GTGTTTCAAC AATACAAyCT CTTTCCGCAC
    301  AAAACCGCCT TGGAAAACGT AATGGAAGGA CCGGTTGCCG TACAgGGCAA
    351  GCCTGCCGCC CAAGCGCGCG AAGAGGCTCT GAAACTGCTG GAAAAAGTCG
    401  GCTTGGGCGA CAAAGTGGAT TTGTATCCCT ACCAGCTTTC CGGCGGTCAG
    451  CAGCAGCGCG TCGGCATTGC CCGCGCATTG GCGATTCAGC CTGAACTGAT
    501  GCTGTTTGAC GAACCGACTT CCGCGCTCGA TCCTGAATTG GTGCAAGATG
    551  TTTTGGATmC CATGAAGGAA TTGGCGCAAG AAGGCTGGAC CATGGTTGTC
    601  GTTACGCATG AAATCAAGTT CGCCTTAGAA GTGGCAACCA CCGwCGTCGT
    651  GATGGACrGC GGCGTTATTG TCGAACAAGG CAGCCCGCAA GATTTGTTCG
    701  ACCACCCCAA ACACGAACGG ACGCGGAGAT TTTTAAGCCA AATCCAATCT
    751  ACCAAGATTT GA
```

This corresponds to the amino acid sequence <SEQ ID 448; ORF 122>:

```
m122.pep
      1  VVMIKIRNIH KTFGENTILR GIDLDVCKGQ VVVILGPSGS GKTTFLRCLN
     51  ALEMPEDGQI EFDNERPLKI DFSKKPSKHD ILALRRKSXM VFQQYNLFPH
    101  KTALENVMEG PVAVQGKPAA QAREEALKLL EKVGLGDKVD LYPYQLSGGQ
    151  QQRVGIARAL AIQPELMLFD EPTSALDPEL VQDVLDXMKE LAQEGWTMVV
    201  VTHEIKFALE VATTXVVMDX GVIVEQGSPQ DLFDHPKHER TRRFLSQIQS
    251  TKI*
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

ORF 122 shows 47.2% identity over a 246 aa overlap with a predicted ORF (ORF 122.ng) from *N. gonorrhoeae:*

```
m122/g122

                      10        20        30        40        50        60
m122.pep      VVMIKIRNIHKTFGENTILRGIDLDVCKGQVVVILGPSGSGKTTFLRCLNALEMPEDGQI
              :::::||::||  :|   |  ::::|||: ||:|:|:||||| ||:|:|||:|:||  : |:|
g122          MALLSIRKLHKQYGSVTAIQSLDLDLEKGEVIVLLGPSGCGKSTLLRCVNGLEPHQGGSI
                      10        20        30        40        50        60

                      70        80        90       100       110       120
m122.pep      EFDNERPLKIDFSKKPSKHDILALRRKSXMVFQQYNLFPHKTALENVMEGPVAVQGKPAA
               :|:   :   | | :  :          |:|  ||||: :|| | |::||:: |||  |:: |
g122          VMDGVGEFGKDVSWQTA-------RQKVGMVFQSNELFAHMTVIENIFLGPVKEQNRDRA
                      70                80        90       100       110

                     130       140       150       160       170       180
m122.pep      QAREEALKLLEKVGLGDKVDLYPYQLSGGQQQRVGIARALAIQPELMLFDEPTSALDPEL
              :|: :|  ||||:|||  |: :  ||  :|||||:||::|:|||  ::||::|:||  |:||||:
g122          EAEAQAGKLLERVGLLDRKNAYPRELSGGQKQRIAIVRALCLNPEVILLDEITAALDPEM
                  120       130       140       150       160       170

                     190       200       210       220       230       240
m122.pep      VQDVLDXMKELAQEGWTMVVVTHEIKFALEVATTXVVMDXGVIVEQGSPQDLFDHPKHER
              |::||: :  |||:||  :|::||||: ||  :||    | ||  |  |||:::|: :|: || ||
g122          VREVLEVVLELAREGMSMLIVTHEMGFARKVADRIVFMDKGGIVESSDPETFFSAPKSER
```

```
                |||||||||||||
a122          TRRFLSQIQSTKIX
                    250
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 451>:

g122-1.seq

```
  1  ATGATTAAAA TCCGCAATAT CCATAAGACC TTTGGCGAAA ACACCATTTT
 51  GCGCGGCATC GATTTGGATG TGGGCAAAGG GCAGGTGGTC GTCATCCTCG
101  GGCCTTCCGG CTCGGGTAAA ACAACATTTC TGCGCTGCCT AAACGCGTTG
151  GAAATGCCCG AAGACGGACA AATCGAGTTC GACAACGCGC GGCCGTTACG
201  CATTGATTTT TCCAAAAAAA CAAGCAAACA CGATATTTTG GCACTGCGCC
251  GCAAGTCCGG AATGGTATTC CAACAATACA ACCTCTTCCC GCATAAAACC
301  GTGTTGGAAA ACGTGATGGA AGGGCCGGTT GCCGTACAGG GCAAGCCTGC
351  CGCCCAAGCG CGCGAAGAGG CTTTGAAACT GCTGGAAAAA GTCGGCTTGG
401  GCGATAAAGT GGATTTGTAT CCCTACCAGC TTTCCGGCGG TCAGCAGCAG
451  CGTGTCGGTA TCGCCCGCGC ACTGGCGATT CAGCCTGAAT TGATGCTGTT
501  TGACGAACCC ACTTCCGCGC TGGACCCCGA GTTGGTGCAA GACGTGTTGG
551  ACGCCATGAA GGAATTGGCG CGGGAAGGTT GGACGATGGT CGTCGTTACC
601  CACGAAATCA AGTTCACGCT GGAAGTTGCC ACGAACGTCG TCGTGATGGA
651  CGGCGGCGTT ATCGTAGAGC AGGGCAGCCC GAAAGAGTTG TTCGACCACC
701  TCAAACACGA ACGGACGCGG AGATTTTTAA GCCAAATCCA ATCTGCCAAG
751  ATTTGA
```

This corresponds to the amino acid sequence <SEQ ID 452; ORF 122-1.ng>:

g122-1.pep

```
  1  MIKIRNIHKT FGENTILRGI DLDVGKGQVV VILGPSGSGK TTFLRCLNAL
 51  EMPEDGQIEF DNARPLRIDF SKKTSKHDIL ALRRKSGMVF QQYNLFPHKT
101  VLENVMEGPV AVQGKPAAQA REEALKLLEK VGLGDKVDLY PYQLSGGQQQ
151  RVGIARALAI QPELMLFDEP TSALDPELVQ DVLDAMKELA REGWTMVVVT
201  HEIKFTLEVA TNVVVMDGGV IVEQGSPKEL FDHLKHERTR RFLSQIQSAK
251  I*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 453>:

m122-1.seq

```
  1  ATGATTAAAA TCCGCAATAT CCATAAGACC TTTGGCGAAA ACACTATTTT
 51  GCGCGGCATC GATTTGGATG TGTGCAAAGG GCAGGTGGTC GTCATCCTCG
101  GGCCTTCCGG CTCAGGCAAA ACGACGTTTC TGCGATGCCT AAACGCGTTG
151  GAAATGCCCG AAGACGGACA AATCGAGTTC GACAACGAGC GACCGCTGAA
201  AATCGATTTT TCTAAAAAAC CAAGCAAACA CGATATTTTG GCACTGCGCC
251  GCAAATCAGG CATGGTGTTT CAACAATACA ACCTCTTTCC GCACAAAACC
301  GCCTTGGAAA ACGTAATGGA AGGACCGGTT GCCGTACAGG GCAAGCCTGC
351  CGCCCAAGCG CGCGAAGAGG CTCTGAAACT GCTGGAAAAA GTCGGCTTGG
401  GCGACAAAGT GGATTTGTAT CCCTACCAGC TTTCCGGCGG TCAGCAGCAG
451  CGCGTCGGCA TTGCCCGCGC ATTGGCGATT CAGCCTGAAC TGATGCTGTT
501  TGACGAACCG ACTTCCGCGC TCGATCCTGA ATTGGTGCAA GATGTTTTGG
551  ATACCATGAA GGAATTGGCG CAAGAAGGCT GGACCATGGT TGTCGTTACG
601  CATGAAATCA AGTTCGCCTT AGAAGTGGCA ACCACCGTCG TCGTGATGGA
651  CGGCGGCGTT ATTGTCGAAC AAGGCAGCCC GCAAGATTTG TTCGACCACC
701  CCAAACACGA ACGGACGCGG AGATTTTTAA GCCAAATCCA ATCTACCAAG
751  ATTTGA
```

This corresponds to the amino acid sequence <SEQ ID 454; ORF 122-1>:

m122-1.pep

```
  1  MIKIRNIHKT FGENTILRGI DLDVCKGQVV VILGPSGSGK TTFLRCLNAL
 51  EMPEDGQIEF DNERPLKIDF SKKPSKHDIL ALRRKSGMVF QQYNLFPHKT
101  ALENVMEGPV AVQGKPAAQA REEALKLLEK VGLGDKVDLY PYQLSGGQQQ
151  RVGIARALAI QPELMLFDEP TSALDPELVQ DVLDTMKELA QEGWTMVVVT
201  HEIKFALEVA TTVVVMDGGV IVEQGSPQDL FDHPKHERTR RFLSQIQSTK
251  I*
```

m122-1/g122-1   94.8% identity in 251 aa overlap

```
                  10        20        30        40        50        60
m122-1.pep  MIKIRNIHKTFGENTILRGIDLDVCKGQVVVILGPSGSGKTTFLRCLNALEMPEDGQIEF
            |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
g122-1      MIKIRNIHKTFGENTILRGIDLDVGKGQVVVILGPSGSGKTTFLRCLNALEMPEDGQIEF
```

```
                  130        140        150        160        170        180

                  190        200        210        220        230        240
a122-1.pep  DVLNAMKELAREGWTMVVVTHEIKFALEVATTVVVMDGGVIVEQGSPKELFDHPKHERTR
            |||::|||||:|||||||||||||||||||||||||||||||||||||::||||||||||:|
m122-1      DVLDTMKELAQEGWTMVVVTHEIKFALEVATTVVVMDGGVIVEQGSPQDLFDHPKHERTR
                  190        200        210        220        230        240

                  250
a122-1.pep  RFLSQIQSTKIX
            ||||||||||||
m122-1      RFLSQIQSTKIX
                  250
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 457>:

```
g125.seq
     1  ATGTCGGGCA ATGCCTCCTC TCCTTCATCT TCCGCCGCCA TCGGGCTGGT
    51  TTGGTTCGGC GCGGCGGTAT CGATTGCCGA AATCAGCACG GGTACGCTGC
   101  TCGCCCCCTT GGGCTGGCAG CGCCGTCTGG CGGCCCTGCT TTTGGGTCAT
   151  GCCGTCGGCG GCGCGCTGTT TTTTGCGGCG GCGTATATCG GCGCACTGAC
   201  CGGACGCAGC TCGATGGAAA GTGTGCGCCT GTCGTTCGGC AAATGCGGTT
   251  CAGTGCTGTT TTCCGTGGCG AATATGCTGC AACTGGCCGG CTGGACGGCG
   301  GTGATGATTT ACGTCGGCGC AacggTCAGC TCCGCTTTGG GCAAAGTGTT
   351  GTGGGACggc gaATCCTTTG TCTGGTGGGC ATTGGCAAAC GGCGCACTGA
   401  TCGTGCTGTG GCTGGTTTTC GGCGCACGCA GAACGGGCGG GCTGAAAACC
   451  GTTTCGATGC TGCTGATGCT GCTTGCCGTG TTGTGGTTGA GCGTCGAAGT
   501  GTTCGCTTCG TCCGGCACAA ACGCCGCGCC CGCCGTTTCA GACGGCATGA
   551  CCTTCGGAAC GGCAGTCGAA CTGTCCGCCG TCATGCCGCT TTCCTGGCTG
   601  CCGCTGGCCG CCGACTACAC GCGCCAAGCA CGCCGCCCGT TTGCGGCAAC
   651  CCTGACGGCA ACGCTCGCCT ATACGCTGAC GGGCTGCTGG ATGTATGCCT
   701  TGGGTTTGGC GGCGGCTCTG TTTACCGGAG AAACCGACGT GGCGAAAATC
   751  CTGTTGGGCG CGGGCTTGGG CATAACGGGC ATTCTGGCAG TCGTCCTCTC
   801  CACCGTTACC ACAACGTTTC TCGATACCTA TTCCGCCGGC GCGAGTGCGA
   851  ACAACATTTC CGCGCGTTTT GCGGAAATAC CCGTCGCTGT CGGCGTTACC
   901  CTGatccgca ccgtgcttgc cgtcatgctg cccgttaccg aatataaaaa
   951  cttcctgctg cttatccgct cggtatttgg gccgatggcg ggtggttttg
  1001  attgccgaCT TTTttgtctt AAAACGGCGT GA
```

This corresponds to the amino acid sequence <SEQ ID 458; ORF 125.ng>:

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 459>:

```
m125.seq
     1  ATGTCGGGCA ATGCCTCCTC TCCTTCATCT TCCTCCGCCA TCGGGCTGAT
    51  TTGGTTCGGC GCGGCGGTAT CGATTGCCGA AATCAGCACG GGTACGCTGC
   101  TTGCGCCTTT GGGCTGGCAG CGCGGTCTGG CGGCTCTACT TTTGGGTCAT
   151  GCCGTCGGCG GCGCGCTGTT TTTTGCGGCG GCGTATATCG GCGCACTGAC
   201  CGGACGCAGC TCGATGGAAA GCGTGCGCCT GTCGTTCGGC AAACGCGGTT
   251  CAGTGCTGTT TTCCGTGGCG AATATGCTGC AACTGGCCGG CTGGACGGCG
   301  GTGATGATTT ACGCCGGCGC AACGGTCAGC TCCGCTTTGG GCAAAGTGTT
   351  GTGGGACGGC GAATCTTTTG TCTGGTGGGC ATTGGCAAAC GGCGCGCTGA
   401  TTGTGCTGTG GCTGGTTTTC GGCGCACGCA AAACAGGCGG GCTGAAAACC
   451  GTTTCGATGC TGCTGATGCT GTTGGCGGTT CTGTGGCTGA GTGCCGAAGT
   501  CTTTTCCACG GCAGGCAGCA CCGCCGCACA GGTTTCAGAC GGCATGAGTT
   551  TCGGAACGGC AGTCGAGCTG TCCGCCGTGA TGCCGCTTTC CTGGCTGCCG
   601  CTTGCCGCCG ACTACACGCG CCACGCGCGC CGCCCGTTTG CGGCAACCCT
   651  GACGGCAACG CTCGCCTACA CGCTGACCGG CTGCTGGATG TATGCCTTGG
   701  GTTTGGCAGC GGCGTTGTTC ACCGGAGAAA CCGACGTGGC AAAAATCCTG
   751  CTGGGCGCAr GTTTGgGTGC GGCAGGCATT TTGGCGGTCG TCCTCTCCAC
   801  CGTTACCACA ACGTTTCTCG ATGCCTATTC CGCCGGCGCG AGTGCGAACA
```

```
 151  GCCGTCGGCG GCGCGCTGTT TTTTGCGGCG GCGTATATCG GCGCACTGAC
 201  CGGACGCAGC TCGATGGAAA GCGTGCGCCT GTCGTTCGGC AAACGCGGTT
 251  CAGTGCTGTT TTCCGTGGCG AATATGCTGC AACTGGCCGG CTGGACGGCG
 301  GTGATGATTT ACGCCGGCGC AACGGTCAGC TCCGCTTTGG GCAAAGTGTT
 351  GTGGGACGGC GAATCTTTTG TCTGGTGGGC ATTGGCAAAC GGCGCGCTGA
 401  TTGTGCTGTG GCTGGTTTTC GGCGCACGCA AAACAGGCGG GCTGAAAACC
 451  GTTTCGATGC TGCTGATGCT GTTGGCGGTT CTGTGGCTGA GTGCCGAAGT
 501  CTTTTCCACG GCAGGCAGCA CCGCCGCACA GGTTTCAGAC GGCATGAGTT
 551  TCGGAACGGC AGTCGAGCTG TCCGCCGTGA TGCCGCTTTC TTGGCTGCCG
 601  CTGGCCGCCG ACTACACGCG CCACGCGCGC CGCCCGTTTG CGGCAACCCT
 651  GACGGCAACG CTCGCCTACA CGCTGACCGG CTGCTGGATG TATGCCTTGG
 701  GTTTGGCAGC GGCGTTGTTC ACCGGAGAAA CCGACGTGGC AAAAATCCTG
 751  CTGGGCGCAG GTTTGGGTGC GGCAGGCATT TTGGCGGTCG TCCTGTCGAC
 801  CGTTACCACC ACTTTTCTCG ATGCCTACTC CGCCGGCGTA AGTGCCAACA
 851  ATATTTCCGC CAAACTTTCG GAAATACCCA TCGCCGTTGC CGTCGCCGTT
 901  GTCGGCACAC TGCTTGCCGT CCTCCTGCCC GTTACCGAAT ATGAAAACTT
 951  CCTGCTGCTT ATCGGCTCGG TATTTGCGCC GATGGCG.GC GGTTTTGATT
1001  GCCGACTTTT TCGTCTTGAA ACGGCGTGA
```

This corresponds to the amino acid sequence <SEQ ID 462; ORF 125.a>:

```
a125.pep
   1  MSGNASSPSS SAAIGLIWFG AAVSIAEIST GTLLAPLGWQ RGLAALLLGH
  51  AVGGALFFAA AYIGALTGRS SMESVRLSFG KRGSVLFSVA NMLQLAGWTA
 101  VMIYAGATVS SALGKVLWDG ESFVWWALAN GALIVLWLVF GARKTGGLKT
 151  VSMLLMLLAV LWLSAEVFST AGSTAAQVSD GMSFGTAVEL SAVMPLSWLP
 201  LAADYTRHAR RPFAATLTAT LAYTLTGCWM YALGLAAALF TGETDVAKIL
 251  LGAGLGAAGI LAVVLSTVTT TFLDAYSAGV SANNISAKLS EIPIAVAVAV
 301  VGTLLAVLLP VTEYENFLLL IGSVFAPMAX GFDCRLFRLE TA*
```

m125/a125 95.6% identity in 342 aa overlap

```
                   10        20        30        40        50        60
m125.pep    MSGNASSPSSSSAIGLIWFGAAVSIAEISTGTLLAPLGWQRGLAALLLGHAVGGALFFAA
            ||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
a125        MSGNASSPSSSAAIGLIWFGAAVSIAEISTGTLLAPLGWQRGLAALLLGHAVGGALFFAA
                   10        20        30        40        50        60

                   70        80        90       100       110       120
m125.pep    AYIGALTGRSSMESVRLSFGKRGSVLFSVANMLQLAGWTAVMIYAGATVSSALGKVLWDG
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a125        AYIGALTGRSSMESVRLSFGKRGSVLFSVANMLQLAGWTAVMIYAGATVSSALGKVLWDG
                   70        80        90       100       110       120

                  130       140       150       160       170       180
m125.pep    ESFVWWALANGALIVLWLVFGARKTGGLKTVSMLLMLLAVLWLSAEVFSTAGSTAAQVSD
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a125        ESFVWWALANGALIVLWLVFGARKTGGLKTVSMLLMLLAVLWLSAEVFSTAGSTAAQVSD
                  130       140       150       160       170       180

                  190       200       210       220       230       240
m125.pep    GMSFGTAVELSAVMPLSWLPLAADYTRHARRPFAATLTATLAYTLTGCWMYALGLAAALF
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a125        GMSFGTAVELSAVMPLSWLPLAADYTRHARRPFAATLTATLAYTLTGCWMYALGLAAALF
                  190       200       210       220       230       240

                  250       260       270       280       290       300
m125.pep    TGETDVAKILLGAXLGAAGILAVVLSTVTTTFLDAYSAGASANNISARFAETPVAVXVTL
            ||||||||||||| |||||||||||||||||||||||||:|||||||::::| |:|| |::
a125        TGETDVAKILLGAGLGAAGILAVVLSTVTTTFLDAYSAGVSANNISAKLSEIPIAVAVAV
                  250       260       270       280       290       300

                  310       320       330       340
m125.pep    IGTVLAVMLPVTEYENFLLLIGSVFAPMAGGFDCRLFRLETAX
            :||:|||:|||||||||||||||||||||| ||||||||||||
a125        VGTLLAVLLPVTEYENFLLLIGSVFAPMAXGFDCRLFRLETAX
```

```
101    DWIKLELIGD DDTLQPDVFQ LVEAAEILIK DGFKVLPYCT EDLIACRRLL
151    DAGCQALMPW AAPIGTGLGA VHAYALNVLR ERLPDTPLII DAGLGLPSQA
201    AQVMEWGFDG VLLNTAVSRS GDPVNMARAF ALAVESGRLA FEAGPVEARD
251    KAQASTPTVG QPFWHSAEY*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 126 shows 95.9% identity over a 269 aa overlap with a predicted ORF (ORF 126.ng) from *N. gonorrhoeae:*

m126/g126

```
                                        10        20        30        40
m126.pep                      HYTKEPIMLTLYGETFPSRLLLGTAAYPTPEILKQSIQTAQ
                              ::||| ||||||||||||||||||||||||||||||::||:
g126      MPSETPKARRRLSDGIASDNHTKESIMLTLYGETFPSRLLLGTAAYPTPEILKQSVRTAR
                  10        20        30        40        50        60

                  50        60        70        80        90       100
m126.pep  PAMITVSLRRAGSGGEAHGQGFWSLLQETGVPVLPNTAGCQSVQEAVTTAQMAREVFETD
          ||||||||||:| ||||||||||||||||||||||||||||||||||||||||||||||||
g126      PAMITVSLRRTGCGGEAHGQGFWSLLQETGVPVLPNTAGCQSVQEAVTTAQMAREVFETD
                  70        80        90       100       110       120

                 110       120       130       140       150       160
m126.pep  WIKLELIGDDDTLQPDVFQLVEAAEILIKDGFKVLPYCTEDLIACRRLLDAGCQALMPWA
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g126      WIKLELIGDDDTLQPDVFQLVEAAEILIKDGFKVLPYCTEDLIACRRLLDAGCQALMPWA
                 130       140       150       160       170       180

                 170       180       190       200       210       220
m126.pep  APIGTGLGAVHAYALNVLRERLPDTPLIIDAGLGLPSQAAQVMEWGFDGVLLNTAVSRSG
          ||||||||||||||||::||||||||||||||||||||||||||||||||||||||||||
g126      APIGTGLGAVHAYALKILRERLPDTPLIIDAGLGLPSQAAQVMEWGFDGVLLNTAVSRSG
                 190       200       210       220       230       240

                 230       240       250       260       270
m126.pep  DPVNMARAFALAVESGRLAFEAGPVEARDKAQASTPTVGQPFWHSAEYX
          ||||||||||||||||||||||||||||| |||||||||||||||||||
g126      DPVNMARAFALAVESGRLAFEAGPVEARTKAQASTPTVGQPFWHSAEYX
                 250       260       270       280
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 467>:

```
a126.seq
      1  TTGTTAATCC ACTATACAAA GGAACCCATT ATGCTCACCC TGTACAGCGA
     51  AACTTTCCCT TCGCGGCTGC TGCTCGGCAC AGCCGCCTAC CCGACCCCTG
    101  AAATCCTCAA ACAATCCGTC CGAACCGCCC GGCCCGCGAT GATTACCGTC
    151  TCGCTGCGCC GCGCGGGATG CGGCGGCGAG GCGCACGGTC AGGGGTTTTG
    201  GTCGCTGCTT CAAGAAACCG GCGTTCCCGT CCTGCCGAAC ACGGCAGGCT
    251  GCCAAAGCGT GCAGGAAGCG GTAACGACGG CGCAAATGGC GCGCGAAGTG
    301  TTTGAAACCG ATTGGATTAA ACTCGAACTC ATCGGCGACG ACGACACCTT
    351  GCAGCCGGAT GTGTTCCAAC TTGTCGAAGC GGCGGAAATC CTGATTAAAG
    401  ACGGCTTCAA AGTGCTGCCT TATTGCACCG AAGACCTGAT TGCCTGCCGC
    451  CGCCTGCTCG ACGCGGGCTG TCAGGCGTTG ATGCCGTGGG CGGCCCCGAT
    501  CGGCACGGGT TTGGGCGCGG TTCACGCCTA CGCGTTGAAC GTCCTGCGCG
    551  AACGCCTGCC CGACACGCCG C[illegible]GATTATCG ACGCGGGCTT GGGTTTGCCC
    601  TCACAGGCGG CACAAGTGAT GGAATGGGGC TTTGACGGCG TGCTTTTGAA
    651  TACTGCCGTT TCCCGCAGCG GCGATCCGGT CAATATGGCA CGCGCCTTCG
    701  CACTCGCCGT CGAATCCGGA CGGCTGGCAT TTGAAGCCGG ACCGGTCGAA
    751  GCACGCGACA AAGCGCAAGC CAGCACGCCG ACAGTCGGAC AACCGTTTTG
    801  GCATTCGGCG GAATATTGA
```

```
 51  AHGQGFWSLL QETGVPVLPN TAGCQSVQEA VTTAQMAREV FETDWIKLEL
101  IGDDDTLQPD VFQLVEAAEI LIKDGFKVLP YCTEDLIACR RLLDAGCQAL
151  MPWAAPIGTG LGAVHAYALK ILRERLPDTP LIIDAGLGLP SQAAQVMEWG
201  FDGVLLNTAV SRSGDPVNMA RAFALAVESG RLAFEAGPVE ARTKAQASTP
251  TVGQPFWHSA EY*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 471>:

m126-1.seq

```
  1  ATGCTCACCC TATACGGCGA AACTTTCCCC TCGCGGCTGC TGCTCGGCAC
 51  GGCTGCCTAC CCGACCCCCG AAATCCTCAA ACAATCCATC CAAACCGCCC
101  AGCCTGCGAT GATTACCGTC TCGCTGCGCC GCGCGGGAAG CGGCGGCGAG
151  GCGCACGGTC AGGGGTTTTG GTCGCTGCTT CAAGAAACCG GCGTTCCCGT
201  CCTGCCGAAC ACGGCAGGCT GCCAAAGCGT GCAGGAAGCG GTAACGACGG
251  CGCAAATGGC GCGCGAAGTG TTTGAAACCG ATTGGATAAA ATTGGAACTC
301  ATCGGAGATG ACGACACCTT GCAGCCGGAT GTGTTCCAGC TTGTCGAAGC
351  GGCGGAAATC CTGATTAAAG ACGGCTTCAA AGTGCTGCCT TATTGCACCG
401  AAGACCTGAT TGCCTGCCGC CGCCTGCTCG ACGCGGGCTG TCAGGCGTTG
451  ATGCCGTGGG CGGCCCCGAT CGGCACGGGT TTGGGCGCGG TTCACGCCTA
501  CGCGTTGAAC GTCCTGCGCG AACGCCTGCC CGACACGCCG CTGATTATCG
551  ACGCGGGCTT GGGTTTGCCC TCACAGGCGG CACAAGTGAT GGAATGGGGC
601  TTTGACGGCG TGCTTTTGAA TACTGCCGTT TCCCGCAGCG GCGATCCGGT
651  CAATATGGCA CGCGCCTTCG CACTCGCCGT CGAATCCGGA CGGCTGGCAT
701  TTGAAGCCGG ACCGGTCGAA GCACGCGACA AAGCGCAAGC CAGCACGCCG
751  ACAGTCGGAC AACCGTTTTG GCATTCGGCG GAATATTGA
```

This corresponds to the amino acid sequence <SEQ ID 5; ORF 126-1>:

m126-1.pep

```
  1  MLTLYGETFP SRLLLGTAAY PTPEILKQSI QTAQPAMITV SLRRAGSGGE
 51  AHGQGFWSLL QETGVPVLPN TAGCQSVQEA VTTAQMAREV FETDWIKLEL
101  IGDDDTLQPD VFQLVEAAEI LIKDGFKVLP YCTEDLIACR RLLDAGCQAL
151  MPWAAPIGTG LGAVHAYALN VLRERLPDTP LIIDAGLGLP SQAAQVMEWG
201  FDGVLLNTAV SRSGDPVNMA RAFALAVESG RLAFEAGPVE ARDKAQASTP
251  TVGQPFWHSA EY*
```

m126-1/g126-1 96.9% identity in 262 aa overlap

```
                     10        20        30        40        50        60
m126-1.pep   MLTLYGETFPSRLLLGTAAYPTPEILKQSIQTAQPAMITVSLRRAGSGGEAHGQGFWSLL
             |||||||||||||||||||||||||||||::||:||||||||||:| |||||||||||||
g126-1       MLTLYGETFPSRLLLGTAAYPTPEILKQSVRTARPAMITVSLRRTGCGGEAHGQGFWSLL
                     10        20        30        40        50        60

                     70        80        90       100       110       120
m126-1.pep   QETGVPVLPNTAGCQSVQEAVTTAQMAREVFETDWIKLELIGDDDTLQPDVFQLVEAAEI
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g126-1       QETGVPVLPNTAGCQSVQEAVTTAQMAREVFETDWIKLELIGDDDTLQPDVFQLVEAAEI
                     70        80        90       100       110       120

                    130       140       150       160       170       180
m126-1.pep   LIKDGFKVLPYCTEDLIACRRLLDAGCQALMPWAAPIGTGLGAVHAYALNVLRERLPDTP
             |||||||||||||||||||||||||||||||||||||||||||||||||::|||||||||
g126-1       LIKDGFKVLPYCTEDLIACRRLLDAGCQALMPWAAPIGTGLGAVHAYALKILRERLPDTP
                    130       140       150       160       170       180

                    190       200       210       220       230       240
m126-1.pep   LIIDAGLGLPSQAAQVMEWGFDGVLLNTAVSRSGDPVNMARAFALAVESGRLAFEAGPVE
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g126-1       LIIDAGLGLPSQAAQVMEWGFDGVLLNTAVSRSGDPVNMARAFALAVESGRLAFEAGPVE
                    190       200       210       220       230       240

                    250       260
m126-1.pep   ARDKAQASTPTVGQPFWHSAEYX
             || ||||||||||||||||||||
g126-1       ARTKAQASTPTVGQPFWHSAEYX
                    250       260
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 473>:

a126-1.seq

```
  1  ATGCTCACCC TGTACAGCGA AACTTTCCCT TCGCGGCTGC TGCTCGGCAC
 51  AGCCGCCTAC CCGACCCCTG AAATCCTCAA ACAATCCGTC CGAACCGCCC
101  GGCCCGCGAT GATTACCGTC TCGCTGCGCC GCGCGGGATG CGGCGGCGAG
```

```
551  CGGTCGAAAT CCCCGTTCCC ATCCATTTGG ATTCGGATGA AGCCGTATGC
601  CGTCTGAAAG CCGTACTCGA GCCCTTGTGC GCGCCCTACA TCCCCGCCAT
651  TCAGCGGTAT TTGGAAAACG TGCAGGCGGA AAAACTGTTT ATCACGCCCG
701  CCGCCAGGCC GCGCGTTACC CGCGTACCGT ACGACGACAA GGCATACCGC
751  ATCATCGTCC GCTTCGCCTC CCCCGTTTCA AAGCGGCTGG AAATCCAACA
801  GGCGGTTATG GACGAATTTT TGCGCGTACA ATACCGCCTG TTAAATCATC
851  CCGCCGgctc cgAAACACTT TAA
```

This corresponds to the amino acid sequence <SEQ ID 476; ORF 127.ng>:

```
g127.pep
  1  MEIWNMLNTW PDAVPIRAEA AESVAAVAAL LLARALLLNI HFRRHPDFGI
 51  ESKRRFLVAS RNITLLLVLF SLAFIWSAQI QTLALSMFAV AAAVVVATKE
101  LIMCLSGSIL RSATQQYSVG DYIEINGLRG RVVDINLLNT LMMQVGPNPL
151  VGQLAGTTVS FPNSLLLSHP VRRDNILGDY VIHTVEIPVP IHLDSDEAVC
201  RLKAVLEPLC APYIPAIQRY LENVQAEKLF ITPAARPRVT RVPYDDKAYR
251  IIVRFASPVS KRLEIQQAVM DEFLRVQYRL LNHPAGSETL *
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 477>:

```
m127.seq
  1  ATGGAAATAT GGAATATGTT GGACACTTGG CTCGGTGCCG TCCCGATACG
 51  TGCGGAGGCG GTCGAATCCG TGGCGGCGGT TGCGGCTTTG CTGCTGGCGC
101  GCGCCCTTCT GTTGAATATC CACTTCAAAC GGCATCCGGA TTTCGGCATC
151  GAAAGCAAGC GGCGGTTTTT GGTTGCCAGC CGCAATATAA CGCTGCTTTT
201  GGTGCTGTTT TCGCTGGCAT TTATCTGGTC GGCGCAAATC CAAACGCTGG
251  CTTTGTCGAT GTTTGCGGTG GCGGCGGCGG TCGTCGTGGC GACGAAGGAA
301  CTGATTATGT GTCTGTCGGG CAGTATTTTA AGGTCTGCCA CCCAGCAATA
351  CTCGGTCGGC GACTATATCG AAATCAACGG CCTGCGCGGG CGCGTGGTCG
401  ACATCAACCT GTTGAACACG CTGATGATGC AGGTCGGTCC GAACCCCTTG
451  GTCGGACAGC TTGCGGGAAC CACCGTTTCT TTCCCCAACA GCCTGTTGTT
501  GAGCCACCCC GTGCGCCGCG ACAATATTTT GGGCGACTAT GTCATCCATA
551  CGGTCGAAAT CCCCGTTCCC ATCCATTTGG ATTCGGATGA AGCCGTATGC
601  CGTCTGAAAG CCGTACTCGA GCCCTTGTGC GCGCCCTACA TCCCCGCCAT
651  CCAACGGsAT TTGGAAAACG TGCAGGCGGA AAAACTGTTT ATCACGCCCG
701  CCGCCAGACC GCGCGTTACC CGCGTGCCGT ACGATGACAA GGCATACCGC
751  ATCATCGTCC GCTTCGCTTC CCCCGTTTCA AAGCGGCTGG AAATCCAACA
801  GGCGGTTATG GACGAATTTT TGCGCGTACA ATACCGCCTG TTAAATCACC
851  CCGCCGGCTC CGAAACACTT TAA
```

This corresponds to the amino acid sequence <SEQ ID 478; ORF 127>:

```
m127.pep
  1  MEIWNMLDTW LGAVPIRAEA VESVAAVAAL LLARALLLNI HFKRHPDFGI
 51  ESKRRFLVAS RNITLLLVLF SLAFIWSAQI QTLALSMFAV AAAVVVATKE
101  LIMCLSGSIL RSATQQYSVG DYIEINGLRG RVVDINLLNT LMMQVGPNPL
151  VGQLAGTTVS FPNSLLLSHP VRRDNILGDY VIHTVEIPVP IHLDSDEAVC
201  RLKAVLEPLC APYIPAIQRX LENVQAEKLF ITPAARPRVT RVPYDDKAYR
251  IIVRFASPVS KRLEIQQAVM DEFLRVQYRL LNHPAGSETL *
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

ORF 127 shows 97.9% identity over a 290 aa overlap with a predicted ORF (ORF 127.ng) from *N. gonorrhoeae:*

```
m127/g127

                    10        20        30        40        50        60
m127.pep    MEIWNMLDTWLGAVPIRAEAVESVAAVAALLLARALLLNIHFKRHPDFGIESKRRFLVAS
            |||||||:||  ||||||||:|||||||||||||||||||||:|||||||||||||||||
g127        MEIWNMLNTWPDAVPIRAEAAESVAAVAALLLARALLLNIHFRRHPDFGIESKRRFLVAS
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m127.pep    RNITLLLVLFSLAFIWSAQIQTLALSMFAVAAAVVVATKELIMCLSGSILRSATQQYSVG
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g127        RNITLLLVLFSLAFIWSAQIQTLALSMFAVAAAVVVATKELIMCLSGSILRSATQQYSVG
```

```
                 130       140       150       160       170       180
m127.pep  DYIEINGLRGRVVDINLLNTLMMQVGPNPLVGQLAGTTVSFPNSLLLSHPVRRDNILGDY
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a127      DYIEINGLRGRVVDINLLNTLMMQVGPNPLVGQLAGTTVSFPNSLLLSHPVRRDNILGDY
                 130       140       150       160       170       180

                 190       200       210       220       230       240
m127.pep  VIHTVEIPVPIHLDSDEAVCRLKAVLEPLCAPYIPAIQRXLENVQAEKLFITPAARPRVT
          ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||:||||
a127      VIHTVEIPVPIHLDSDEAVCRLKAVLEPLCAPYIPAIQRHLENVQAEKLFITPAAKPRVT
                 190       200       210       220       230       240

                 250       260       270       280       290
m127.pep  RVPYDDKAYRIIVRFASPVSKRLEIQQAVMDEFLRVQYRLLNHPAGSETLX
          |||||||||||||||||||||||||||||||||||||||||||:|||||||
a127      RVPYDDKAYRIIVRFASPVSKRLEIQQAVMDEFLRVQYRLLNYPAGSETLX
                 250       260       270       280       290
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 481>:

```
g128.seq
     1  atgattgaca acgCActgct ccacttgggc gaagaacccC GTTTTaatca
    51  aatccaaacc gaagACAtca AACCCGCCGT CCAAACCGCC ATCGCCGAAG
   101  CGCGCGGACA AATCGCCGCC GTCAAAGCGC AAACGCACAC CGGCTGGGCG
   151  AACACCGTCG AGCGTCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
   201  GGGCGTCGTG TCCCATCTCA ACTCCGTCGT CGACACGCCC GAACTGCGCG
   251  CCGTCTATAA CGAACTGATG CCTGAAATCA CCGTCTTCTT CACCGAAATC
   301  GGACAAGACA TCGAACTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
   351  CGAATTTGCA ACGCTTTCCC CCGCACAAAA AACCAAGCTC GATCACGACC
   401  TGCGCGATTT CGTATTGAGC GGCGCGGAAC TGCCGCCCGA ACGGCAGGCA
   451  GAACTGGCAA AACTGCAAAC CGAAGGCGCG CAACTTTCCG CCAAATTCTC
   501  CCAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTTAC TTTGACGATG
   551  CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCC
   601  GCCGCGCAAA GCGAAGGCAA AACAGGTTAC AAAATCGGCT TGCAGATTCC
   651  GCACTACCTT GCCGTTATCC AATACGCCGG CAACCGCGAA CTGCGCGAAC
   701  AAATCTACCG CGCCTACGTT ACCCGTGCCA GCGAACTTTC AAACGACGGC
   751  AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCATTGAA
   801  AACCGccaaa cTGCTCGGCT TTAAAAATTA CGCCGAATTG TCGCTGGCAA
   851  CCAAAATGGC GGACACGCCC GAACAGGTTT TAAACTTCCT GCACGACCTC
   901  GCCCGCCGCG CCAAACCCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
   951  CTTCGCCCGC GAACACCTCG GTCTCGCCGA CCCGCAGCCG TGGGACTTGA
  1001  GCTACGCCGG CGAAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
  1051  GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTTCTGGCAG GCCTGTTCGC
  1101  CCAAATCAAA AAACTCTACG GCATCGGATT CGCCGAAAAA ACCGTTCCCG
  1151  TCTGGCACAA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGGCAAAACC
  1201  ATCGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
  1251  CGCGTGGATG AACGACtaca AAGGCCGCCG CCGCTTTGCC GACGgcacGC
  1301  TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCGCCCC GCCCGTCGGC
  1351  GGCAAAGAAG CGCGTTTAAG CCACGACGAA ATCCTCACCC TCTTCCACGA
  1401  AacCGGCCAC GGACTGCACC ACCTGCTTAC CCAAGTGGAC GAACTGGGCG
  1451  TGTCCGGCAT CAAcggcgtA GAATGGGACG CGGTCGAACT GCCCAGCCAG
  1501  TTTATGGAAA ACTTCGTTTG GGAATACAAT GTATTGGCAC AAATGTCCGC
  1551  CCACGAAGAA AccgGCGAGC CCCTGCCGAA AGAACTCTTC GACAAAATGC
  1601  TcgcCGCCAA AAACTTCCAG CGCGGTATGT TCCTCGTCCG GCAAATGGAG
  1651  TTCGCCCTCT TCGATATGAT GATTTACAGT GAAAGCGACG AATGCCGTCT
  1701  GAAAAACTGG CAGCAGGTTT TAGACAGCGT GCGCAAAGAA GTcGCCGTCA
  1751  TCCAACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCacatctTC
  1801  GCcggcGGCT ATTCCGCAGG CTATTACAGC TACGCATGGG CCGAAGTCCt
  1851  cAGCACCGAT GCCTACGCCG CCTTTGAAGA AAGcGACGac gtcGCCGCCA
  1901  CAGGCAAACG CTTCTGGCAA GAAAtccttg ccgtcggcgg ctCCCGCAGC
  1951  gcgGCGGAAT CCTTCAAAGC CTTCCGCGGA CGCGAACCGA GCATAGACGC
```

```
301  GKRFWQEILA VGXSRSGAES FKAFRGREPS IDALLRHSGF DNAV*
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

ORF 128 shows 91.7% identity over a 475 aa overlap with a predicted ORF (ORF 128.ng) from *N. gonorrhoeae*:

```
m128/g128

                    10        20        30        40        50        60
g128.pep    MIDNALLHLGEEPRFNQIQTEDIKPAVQTAIAEARGQIAAVKAQTHTGWANTVERLTGIT
            | |||||||||||||:||:|||||||:|||||||| ||||:||||||||||||| |||||
m128        MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIKAQTHTGWANTVEPLTGIT
                    10        20        30        40        50        60

                    70        80        90       100       110       120
g128.pep    ERVGRIWGVVSHLNSVVDTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFA
            |||||||||||||| |:|||||||||||||||||||||||||||||||||||||||||| 
m128        ERVGRIWGVVSHLNCVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
                    70        80        90       100       110       120

                   130       140       150       160       170       180
g128.pep    TLSPAQKTKLDHDLRDFVLSGAELPPERQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY
            ||||||||||:|
m128        TLSPAQKTKLNH
                   130
             //
                                              340       350       360
g128.pep                                  YAGEKLREAKYAFSETEVKKYFPVGKVLAG
                                          ||:||||||||||||| |||||||| || |
m128                                      YASEKLREAKYAFSETXVKKYFPVGXVLNG
                                                  10        20        30

                 370       380       390       400       410       420
g128.pep    LFAQIKKLYGIGFAEKTVPVWHKDVRYFELQQNGKTIGGVYMDLYAREGKRGGAWMNDYK
            |||| ||||||||:|||||||||||||| ||||||::|||||||||||||||||||||||
m128        LFAQXKKLYGIGFTEKTVPVWHKDVRYXELQQNGEXIGGVYMDLYAREGKRGGAWMNDYK
                    40        50        60        70        80        90

                 430       440       450       460       470       480
g128.pep    GRRRFADGTLQLPTAYLVCNFAPPVGGKEARLSHDEILTLFHETGHGLHHLLTQVDELGV
            |||||:|||||||||||||||||||||:|||||||||| |||||||||||||||||||||
m128        GRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILILFHETGHGLHHLLTQVDELGV
                   100       110       120       130       140       150

                 490       500       510       520       530       540
g128.pep    SGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGEPLPKELFDKMLAAKNFQRGMF
            |||||| ||||||||||||||||||||||| ||||||| |||||| || ||||||| |||
m128        SGINGVXWDAVELPSQFMENFVWEYNVLAQXSAHEETGVPLPKELXDKXLAAKNFQXGMF
                   160       170       180       190       200       210

                 550       560       570       580       590       600
g128.pep    LVRQMEFALFDMMIYSESDECRLKNWQQVLDSVRKEVAVIQPPEYNRFANSFGHIFAGGY
             ||| ||||||||||||:|| ||||||||||||||:|||||||||||||| |||||||||
m128        XVRQXEFALFDMMIYSEDDEGRLKNWQQVLDSVRKKVAVIQPPEYNRFALSFGHIFAGGY
                   220       230       240       250       260       270

                 610       620       630       640       650       660
g128.pep    SAGYYSYAWAEVLSTDAYAAFEESDDVAATGKRFWQEILAVGGSRSAAESFKAFRGREPS
            ||: ||||||||||:|||||||||||||||||||||||||||| |||:||||||||||||
m128        SAAXYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGXSRSGAESFKAFRGREPS
                   280       290       300       310       320       330
```

```
551  FALFDMMIYS EDDEGRLKNW QQVLDSVRKE VAVVRPPEYN RFANSFGHIF
601  AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRFWQ EILAVGGSRS
651  AAESFKAFRG REPSIDALLR HSGFDNAA*
```

m128/a128 66.0% identity in 677 aa overlap

```
                    10        20        30        40        50        60
m128.pep    MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIKAQTHTGWANTVEPLTGIT
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a128        MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIKAQTHTGWANTVEPLTGIT
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m128.pep    ERVGRIWGVVSHLNCVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
            |||||||||||||| |:||||||:||||||||||||||||||||||||||||||||||||
a128        ERVGRIWGVVSHLNSVTDTPELRAAYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
                    70        80        90       100       110       120

                   130
m128.pep    TLSPAQKTKLNH------------------------------------------------
            ||| ||||||||
a128        TLSHAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY
                   130       140       150       160       170       180

m128.pep    ------------------------------------------------------------

a128        FDDAAPLAGIPEDALAMFAAAAQSEGKTGYKIGLQIPHYLAVIQYADNRKLREQIYRAYV
                   190       200       210       220       230       240

m128.pep    ------------------------------------------------------------

a128        TRASELSDDGKFDNTANIDRTLENALQTAKLLGFKNYAELSLATKMADTPEQVLNFLHDL
                   250       260       270       280       290       300

                                                   140       150
m128.pep    ----------------------------------YASEKLREAKYAFSETXVKKYFPVGX
                                              ||:||||||||||||| |||||||
a128        ARRAKPYAEKDLAEVKAFARESLGLADLQPWDLGYAGEKLREAKYAFSETEVKKYFPVGK
                   310       320       330       340       350       360

           160       170       180       190       200       210
m128.pep    VLNGLFAQXKKLYGIGFTEKTVPVWHKDVRYXELQQNGEXIGGVYMDLYAREGKRGGAWM
            |||||||| |||||||||||||||||||||| ||||||:|||||||||||||||||||||
a128        VLNGLFAQIKKLYGIGFTEKTVPVWHKDVRYFELQQNGETIGGVYMDLYAREGKRGGAWM
                   370       380       390       400       410       420

           220       230       240       250       260       270
m128.pep    NDYKGRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILILFHETGHGLHHLLTQVD
            |||||||||||||||||||||||||:|||||:|||||||||| ||||||||||||||||
a128        NDYKGRRRFSDGTLQLPTAYLVCNFTPPVGGKEARLSHDEILTLFHETGHGLHHLLTQVD
                   430       440       450       460       470       480

           280       290       300       310       320       330
m128.pep    ELGVSGINGVXWDAVELPSQFMENFVWEYNVLAQXSAHEETGVPLPKELXDKXLAAKNFQ
            |||||||||| ||||||||||||||||||||||| |||||||||||||| || |||||||
a128        ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ
                   490       500       510       520       530       540

           340       350       360       370       380       390
m128.pep    XGMFXVRQXEFALFDMMIYSEDDEGRLKNWQQVLDSVRKKVAVIQPPEYNRFALSFGHIF
             ||| ||| ||||||||||||||||||||||||||||:|||::|||||||| ||||||
a128        RGMFLVRQMEFALFDMMIYSEDDEGRLKNWQQVLDSVRKEVAVVRPPEYNRFANSFGHIF
                   550       560       570       580       590       600
```

```
 201  GGGCGTGGTG TCGCACCTCA ACTCCGTCGC CGACACGCCC GAACTGCGCG
 251  CCGTCTATAA CGAACTGATG CCCGAAATCA CCGTCTTCTT CACCGAAATC
 301  GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
 351  CGAATTCGAC ACCCTCTCCC CCGCACAAAA AACCAAACTC AACCACGATC
 401  TGCGCGATTT CGTCCTCAGC GGCGCGGAAC TGCCGCCCGA ACAGCAGGCA
 451  GAACTGGCAA AACTGCAAAC CGAAGGCGCG CAACTTTCCG CCAAATTCTC
 501  CCAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTTAC TTTGACGATG
 551  CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCC
 601  GCCGCGCAAA GCGAAAGCAA AACAGGCTAC AAAATCGGCT TGCAGATTCC
 651  ACACTACCTC GCCGTCATCC AATACGCCGA CAACCGCGAA CTGCGCGAAC
 701  AAATCTACCG CGCCTACGTT ACCCGCGCCA GCGAACTTTC AGACGACGGC
 751  AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGCAA ACGCCCTGCA
 801  AACCGCCAAA CTGCTCGGCT TCAAAAACTA CGCCGAATTG TCGCTGGCAA
 851  CCAAAATGGC GGACACGCCC GAACAAGTTT TAAACTTCCT GCACGACCTC
 901  GCCCGCCGCG CCAAACCCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
 951  CTTCGCCCGC GAAAGCCTGA ACCTCGCCGA TTTGCAACCG TGGGACTTGG
1001  GCTACGCCAG CGAAAAACTG CGCGAAGCCA AATACGCGTT CAGCGAAACC
1051  GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTATTAAACG GACTGTTCGC
1101  CCAAATCAAA AAACTCTACG GCATCGGATT TACCGAAAAA ACCGTCCCCG
1151  TCTGGCACAA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGGCGAAACC
1201  ATAGGCGGCG TTTATATGGA TTTGTACGCA CGCC[illegible]AGGCA AACGCGGCGG
1251  CGCGTGGATG AACGACTACA AAGGCCGCCG CCGTTTTTCA GACGGCACGC
1301  TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCGCCCC [illegible]CGTCGGC
1351  GGCAGGGAAG CCCGCCTGAG CCACGACGAA ATCCTCATCC [illegible]CACGA
1401  AACCGGACAC GGGCTGCACC ACCTGCTTAC CCAAGTGGAC GAACTGGGCG
1451  TATCCGGCAT CAACGGCGTA GAATGGGACG CGGTCGAACT GCCCAGCCAG
1501  TTTATGGAAA ATTTCGTTTG GGAATACAAT GTCTTGGCAC AAATGTCAGC
1551  CCACGAAGAA ACCGGCGTTC CCCTGCCGAA AGAACTCTTC GACAAAATGC
1601  TCGCCGCCAA AAACTTCCAA CGCGGCATGT TCCTCGTCCG GCAAATGGAG
1651  TTCGCCCTCT TTGATATGAT GATTTACAGC GAAGACGACG AAGGCCGTCT
1701  GAAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAAAA GTCGCCGTCA
1751  TCCAGCCGCC CGAATACAAC CGCTTCGCCT TGAGCTTCGG CCACATCTTC
1801  GCAGGCGGCT ATTCCGCAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
1851  GAGCGCGGAC GCATACGCCG CCTTTGAAGA AAGCGACGAT GTCGCCGCCA
1901  CAGGCAAACG CTTTTGGCAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC
1951  GCGGCAGAAT CCTTCAAAGC CTTCCGCGGC CGCGAACCGA GCATAGACGC
2001  ACTCTTGCGC CACAGCGGTT TCGACAACGC GGTCTGA
```

This corresponds to the amino acid sequence <SEQ ID 490; ORF 128-1>:

m128-1.pep.

```
  1  MTDNALLHLG EEPRFDQIKT EDIKPALQTA IAEAREQIAA IKAQTHTGWA
 51  NTVEPLTGIT ERVGRIWGVV SHLNSVADTP ELRAVYNELM PEITVFFTEI
101  GQDIELYNRF KTIKNSPEFD TLSPAQKTKL NHDLRDFVLS GAELPPEQQA
151  ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201  AAQSESKTGY KIGLQIPHYL AVIQYADNRE LREQIYRAYV TRASELSDDG
251  KFDNTANIDR TLANALQTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
301  ARRAKPYAEK DLAEVKAFAR ESLNLADLQP WDLGYASEKL REAKYAFSET
351  EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVWHKDVR YFELQQNGET
401  IGGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTLQLPTAY LVCNFAPPVG
451  GREARLSHDE ILILFHETGH GLHHLLTQVD ELGVSGINGV EWDAVELPSQ
501  FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME
551  FALFDMMIYS EDDEGRLKNW QQVLDSVRKK VAVIQPPEYN RFALSFGHIF
601  AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRFWQ EILAVGGSRS
651  AAESFKAFRG REPSIDALLR HSGFDNAV*
```

m128-1/g128-1 94.5% identity in 491 aa overlap

```
                     10        20        30        40        50        60
g128-1.pep   MIDNALLHLGEEPRFNQIKTEDIKPAVQTAIAEARGQIAAVKAQTHTGWANTVERLTGIT
             | |||||||||||||:||||||||||:|||||||| ||||:||||||||||||| |||||
m128-1       MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIKAQTHTGWANTVEPLTGIT
                     10        20        30        40        50        60

                     70        80        90       100       110       120
g128-1.pep   ERVGRIWGVVSHLNSVVDTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFA
             ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||| 
m128-1       ERVGRIWGVVSHLNSVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
                     70        80        90       100       110       120
```

```
1251  CGCGTGGATG AACGACTACA AAGGCCGCCG CCGTTTTTCA GACGGCACGC
1301  TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCACCCC GCCCGTCGGC
1351  GGCAAAGAAG CCCGCTTGAG CCATGACGAA ATCCTCACCC TCTTCCACGA
1401  AACCGGACAC GGCCTGCACC ACCTGCTTAC CCAAGTCGAC GAACTGGGCG
1451  TATCCGGCAT CAACGGCGTA GAATGGGACG CAGTCGAACT GCCCAGTCAG
1501  TTTATGGAAA ATTTCGTTTG GGAATACAAT GTCTTGGCGC AAATGTCCGC
1551  CCACGAAGAA ACCGGCGTTC CCCTGCCGAA AGAACTCTTC GACAAAATGC
1601  TCGCCGCCAA AAACTTCCAA CGCGGAATGT TCCTCGTCCG CCAAATGGAG
1651  TTCGCCCTCT TTGATATGAT GATTTACAGC GAAGACGACG AAGGCCGTCT
1701  GAAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAGAA GTCGCCGTCG
1751  TCCGACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCACATCTTC
1801  GCAGGCGGCT ATTCCGCAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
1851  GAGCGCGGAC GCATACGCCG CCTTTGAAGA AAGCGACGAT GTCGCCGCCA
1901  CAGGCAAACG CTTTTGGCAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC
1951  GCGGCAGAAT CCTTCAAAGC CTTCCGCGGA CGCGAACCGA GCATAGACGC
2001  ACTCTTGCGC CACAGCGGCT TCGACAACGC GGCTTGA
```

This corresponds to the amino acid sequence <SEQ ID 492; ORF 128-1.a>:

a128-1.pep

```
  1  MTDNALLHLG EEPRFDQIKT EDIKPALQTA IAEAREQIAA IKAQTHTGWA
 51  NTVEPLTGIT ERVGRIWGVV SHLNSVTDTP ELRAAYNELM PEITVFFTEI
101  GQDIELYNRF KTIKNSPEFD TLSHAQKTKL NHDLRDFVLS GAELPPEQQA
151  ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201  AAQSEGKTGY KIGLQIPHYL AVIQYADNRK LREQIYRAYV TRASELSDDG
251  KFDNTANIDR TLENALQTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
301  ARRAKPYAEK DLAEVKAFAR ESLGLADLQP WDLGYAGEKL REAKYAFSET
351  EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVWHKDVR YFELQQNGET
401  IGGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTLQLPTAY LVCNFTPPVG
451  GKEARLSHDE ILTLFHETGH GLHHLLTQVD ELGVSGINGV EWDAVELPSQ
501  FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME
551  FALFDMMIYS EDDEGRLKNW QQVLDSVRKE VAVVRPPEYN RFANSFGHIF
601  AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRFWQ EILAVGGSRS
651  AAESFKAFRG REPSIDALLR HSGFDNAA*
```

m128-1/a128-1 97.8% identity in 677 aa overlap

```
                    10        20        30        40        50        60
a128-1.pep  MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIKAQTHTGWANTVEPLTGIT
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m128-1      MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIKAQTHTGWANTVEPLTGIT
                    10        20        30        40        50        60

                    70        80        90       100       110       120
a128-1.pep  ERVGRIWGVVSHLNSVTDTPELRAAYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
            ||||||||||||||||:||||||:||||||||||||||||||||||||||||||||||||
m128-1      ERVGRIWGVVSHLNSVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
                    70        80        90       100       110       120

                   130       140       150       160       170       180
a128-1.pep  TLSHAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY
            ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m128-1      TLSPAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY
                   130       140       150       160       170       180

                   190       200       210       220       230       240
a128-1.pep  FDDAAPLAGIPEDALAMFAAAAQSEGKTGYKIGLQIPHYLAVIQYADNRKLREQIYRAYV
            ||||||||||||||||||||||||:|||||||||||||||||||||||||:|||||||||
m128-1      FDDAAPLAGIPEDALAMFAAAAQSESKTGYKIGLQIPHYLAVIQYADNRELREQIYRAYV
                   190       200       210       220       230       240

                   250       260       270       280       290       300
a128-1.pep  TRASELSDDGKFDNTANIDRTLENALQTAKLLGFKNYAELSLATKMADTPEQVLNFLHDL
            |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
m128-1      TRASELSDDGKFDNTANIDRTLANALQTAKLLGFKNYAELSLATKMADTPEQVLNFLHDL
                   250       260       270       280       290       300

                   310       320       330       340       350       360
a128-1.pep  ARRAKPYAEKDLAEVKAFARESLGLADLQPWDLGYAGEKLREAKYAFSETEVKKYFPVGK
            |||||||||||||||||||||||:|||||||||||:||||||||||||||||||||||||
```

```
Query: 363 NGLFAQIKKLYGIGFTE-KTVPVWHKDVRYFEL-QQNGETIGGVYMDLYAREGKRGGAWM 420
           +GLF  IK+++ I    E K V  WHKDVR+F+L  +N +   G  Y+DLYARE KRGGAWM
Sbjct: 365 SGLFELIKRIFNIRAVERKGVDTWHKDVRFFDLIDENDQLRGSFYLDLYAREHKRGGAWM 424

Query: 421 NDYKGRRRFSDGTLQLPTAYLVCNFTPPVGGKEARLSHDEIXXXXXXXXXXXXXXXXQVD 480
           +D  GR+R  DG+++ P AYL CNF  P+G K A  +H+E+                Q+D
Sbjct: 425 DDCIGRKRKLDGSIETPVAYLTCNFNAPIGNKPALFTHNEVTTLFHEFGHGIHHMLTQID 484

Query: 481 ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ 540
              V+GINGV WDAVELPSQFMEN+ WE   LA +S H ETG PLPKE   ++L AKNFQ
Sbjct: 485 VSDVAGINGVPWDAVELPSQFMENWCWEEEALAFISGHYETGEPLPKEKLTQLLKAKNFQ 544

Query: 541 RGMFLVRQMEFALFDMMIYSEDDEGRLKNWQQVLDSVRKEVAVVRPPEYNRFANSFGHIF 600
             MF++RQ+EF +FD  ++   D  +         L SV+ +VAV++  ++ R  +SF HIF
Sbjct: 545 AAMFILRQLEFGIFDFRLHHTFDAEKTNQILDTLKSVKSQVAVIKGVDWARAPHSFSHIF 604

Query: 601 XXXXXXXXXXXXWAEVLSADAYAAFEESDDV-AATGKRFWQEILAVGGSRSAAESFKAFR 659
                       WAEVLSADAY+ FEE       TGK F  EIL  GGS    E FK FR
Sbjct: 605 AGGYAAGYYSYLWAEVLSADAYSRFEEEGIFNPITGKSFLDEILTRGGSEEPMELFKRFR 664

Query: 660 GREPSIDALLRHSGFDN 676
           GREP +DALLRH G  N
Sbjct: 665 GREPQLDALLRHKGIMN 681
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 493>:

```
g129.seq
      1  ATGCTTTCAC CTCCTCGGCG TAAAACGGCG GCACATCAAT CAAGCCGTCT
     51  TTCATTTGCG TGCGGAAAAA ATGCGGCGTG TTGCCGTGAT CAAAATCAAT
    101  ATCGTGCAGC ATCCAGCCCA AATCGCGGTT TGCCTCGCTT TCCGATAACG
    151  CCGACGGCGG CAGCGGTTCA CCCTTATCCG CGCTTTCGCC ATTTGCCCTT
    201  TCAGGCTGCG GGCATAGGGG CGGAACAGGC GGCGGTCGAA TCCTGTTTCA
    251  TCCGGACAAA CGCGTTGGCA GTCGGAAAAT CCGGCCGGCC GTGTCAAATA
    301  ATGCGTTACT TTGGCCGGGT CTTGTCCTTT GTAAGCGGCG GTCTTTTTTT
    351  GCGCGCCATC CGCATCTGTT TGGGCGCATG GCAAACGGCG GCTGCCGTAC
    401  AATCAAAATG TTTGGCGATT TCATGCAGAC AGGCATCCGG ATGCCGCCCG
    451  ACATATCGAG CCGGTTTTTG CCTATCCGAT TTGGCGGCAT TTAGGCCGGT
    501  AACTTGA
```

This corresponds to the amino acid sequence <SEQ ID 494; ORF 129.ng>:

```
g129.pep
      1  MLSPPRRKTA AHQSSRLSFA CGKNAACCRD QNQYRAASSP NRGLPRFPIT
     51  PTAAAVHPYP RFRHLPFQAA GIGAEQAAVE SCFIRTNALA VGKSGRPCQI
    101  MRYFGRVLSF VSGGLFLRAI RICLGAWQTA AAVQSKCLAI SCRQASGCRP
    151  TYRAGFCLSD LAAFRPVT*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 495>:

```
m129.seq  (partial)
      1  ..TATCTGCGCT TTCACTATTT GCCCTTTCAG GCTGCGGGCA TAGGGACGGA
     51    ACAGGTAGCG GTCAAATCCT GTTTCATCCA AATAAACACG TTGGTAGTCG
    101    GAAAATTCGG CCGGCTGTGT CAAATAATGC GTTACTTTGG CCGGGTCTTG
    151    TTCTTTGTAA GTGGTGGTCT TTTTTTGCGC GTTATCCCCA TCTGTTTGAG
    201    TGCATAGCAA ATGGTGGCTG CCGTACAATC AAAATGTTTG GCGATTTCAT
    251    GCAGATAGGC ATCCGGGTGT TGCCCAACAT ATTGAGCCGG TTTTTGCCTA
    301    TCCGATTTGA CGGCATTTAG ACCGGTAACT TGA
```

This corresponds to the amino acid sequence <SEQ ID 496; ORF 129>:

```
m129.pep  (partial)
      1  ..YLRFHYLPFQ AAGIGTEQVA VKSCFIQINT LVVGKFGRLC QIMRYFGRVL
     51    FFVSGGLFLR VIPICLSAXQ MVAAVQSKCL AISCRXASGC CPTYXAGFCL
    101    SDLTAFRPVT *
```

Computer analysis of this amino acid sequence gave the following results:

```
251  AATGCCACGC GGCGGACAGC AATGTGCCGA ACGCTCCGAA ACTGGAACAC
301  AACGGCGACT GGGCGCCGCG TATCGCGCAA GGCTTCGATA CCTTGTTCCA
351  ACACGCGCTG AACGGCTTTA ACGCCATGCC TGCCAAAGGC GGTGCGGCAG
401  ACCTGACCGA TCAGGAACTC AAACGGGCGA TTACCTACAT GGCGAATAAA
451  AGCGGCGGTT CTTTCCCGAA TCCTGATGAG GCTGCGCCTG CCGACAATGC
501  CGCTTCAGGA ACAGCTTCTG CTCCTGCCGA TAGTGCAGCT CCGGCAGAAG
551  CGAAGGCAGA AGACAAGGGT GCGGCAGCCC CTGCGGTCGG CGTTGACGGT
601  AAAAAAGTCT TCGAAGCAAC CTGTCAGGTG TGCCACGGCG GTTCGATTCC
651  CGGTATTCCC GGCATAGGCA AAAAAGACGA TTGGGCACCG CGTATCAAAA
701  AAGGCAAAGA AACCTTGCAC AAACATGCCC TTGAAGGCTT TAACGCGATG
751  CCGGCCAAAG GCGGCAATGC AGGTTTGAGC GATGACGAAG TCAAAGCGGC
801  TGTTGACTAT ATGGCAAACC AATCCGGTGC AAAATTCTAA
```

This corresponds to the amino acid sequence <SEQ ID 500; ORF 130.ng>:

```
g130.pep
  1  MKQLRDNKAQ GSALFTLVSG IVIVIAVLYF LIKLAGSGSF GDVDATTEAA
 51  TQTRIQPVGQ LTMGDGIPVG ERQGEQIFGK ICIQCHAADS NVPNAPKLEH
101  NGDWAPRIAQ GFDTLFQHAL NGFNAMPAKG GAADLTDQEL KRAITYMANK
151  SGGSFPNPDE AAPADNAASG TASAPADSAA PAEAKAEDKG AAAPAVGVDG
201  KKVFEATCQV CHGGSIPGIP GIGKKDDWAP RIKKGKETLH KHALEGFNAM
251  PAKGGNAGLS DDEVKAAVDY MANQSGAKF*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 501>:

```
m130.seq  (partial)
  1  ..GGCGAACAGA TTTTCGGCAA AATCTGTATC CAATGCCACG CGGCGGACAG
 51    CAATGTGCCG AACGCTCCGA AACTGGAACA CAACGGCGAT TrGGCACCGC
101    GTATCGGgCAA GGCTTCGATA CCTTGTTCCA ACACGCGCTG AACGGCTTTA
151    ACGCCATGCC TGCAAAAGGC GGTGCGGCAG ACCTGACCGA TCAGGAACTT
201    AAACGGGCGA TTACTTACAT GGCGAACAAA AGCGGCGGTT CTTTCCCGAA
251    TCCTGATGAG GCTGCGCCTG CCGACAATGC CGCTTCAGGA ACAGCTTCTG
301    CTCCTGCCGA TAGTGCAGCT CCGGCAGAAG CGAAGGCAGA AGACAAGGGT
351    GCGGCAcCCC TGCGGTCGGC GTTGACGGTA AAAAAGTCTT CGAAGCAACC
401    TGTCAGGTGT GCCACGGCGG TTCGATTCCC GGTATTCCCG GCATAGGCAA
451    AAAAGACGAT TGGGCACCGC GTATCAAAAA AGGCAAAGAA ACCTTGCACA
501    AACACGCCCT TGAAGGCTTT AACGCGATGC CTGCCAAArG CGgCAATGCA
551    GGTTTGAGCG ATGACGAAgT CAAAGCGGCT GTTGACTATA TGGCAAACCA
601    ATCCGGTGCA AAATTCTAA
```

This corresponds to the amino acid sequence <SEQ ID 502; ORF 130>:

```
m130.pep  (partial)
  1  ..GEQIFGKICI QCHAADSNVP NAPKLEHNGD XAPRIQGFDT LFQHALNGFN
 51    AMPAKGGAAD LTDQELKRAI TYMANKSGGS FPNPDEAAPA DNAASGTASA
101    PADSAAPAEA KAEDKGAAPA VGVDGKKVFE ATCQVCHGGS IPGIPGIGKK
151    DDWAPRIKKG KETLHKHALE GFNAMPAKXG NAGLSDDEVK AAVDYMANQS
201    GAKF*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 130 shows 98.1% identity over a 206 aa overlap with a predicted ORF (ORF 130.ng) from *N. gonorrhoeae:*

```
m130/g130

                                                   10        20        30
m130.pep                                 GEQIFGKICIQCHAADSNVPNAPKLEHNGD
                                         ||||||||||||||||||||||||||||||
g130          DATTEAATQTRIQPVGQLTMGDGIPVGERQGEQIFGKICIQCHAADSNVPNAPKLEHNGD
                    50        60        70        80        90       100

                      40        50        60        70        80        89
m130.pep      XAPRI-QGFDTLFQHALNGFNAMPAKGGAADLTDQELKRAITYMANKSGGSFPNPDEAAP
               |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
g130          WAPRIAQGFDTLFQHALNGFNAMPAKGGAADLTDQELKRAITYMANKSGGSFPNPDEAAP
                   110       120       130       140       150       160
```

```
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
a130         KKDDWAPRIKKGKETLHKHALEGFNAMPAKGGNAGLSDDEVKAAVDYMANQSGAKFX
                   230       240       250       260       270       280
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 505>:

```
g132.seq
      1  ATGGAAGCCT TCAAAACCCT AATTTGGATT ATTAATATTA TTTCCGCTTT
     51  GGCCGTCATC GTGTTAGTAT TGCTCCAACA CGGCAAAGGC GCGGATGCCG
    101  GCGCGACCTT CGGATCGGGA AGCGGCAGCG CGCAAGGCGT ATTCGGCTCT
    151  GCCGGCAACG CCAACTTcct CAgccGCTCG AccGccGTTG CAGCAACAtt
    201  tttctttGCa acctgcAtgg gctatggTgt atattcacac CCACACGACA
    251  AAACACGGTT TGGACTtcag caacataCGA CAGACTCAGC AagcACCCAA
    301  ACCcgtAAGC AATACCGAAC CTTCTGCCCC TGTTCCTCAG CAGCAGAAAT
    351  AACagtTTTT CAAATgccga caTGgtga
```

This corresponds to the amino acid sequence <SEQ ID 506; ORF 132.ng>:

```
g132.pep
      1  MEAFKTLIWI INIISALAVI VLVLLQHGKG ADAGATFGSG SGSAQGVFGS
     51  AGNANFLSRS TAVAATFFFA TCMGYGVYSH PHDKTRFGLQ QHTTDSASTQ
    101  TRKQYRTFCP CSSAAEITVF QMPTW*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 507>:

```
m132.seq  (partial)
      1  ATGGAACCCT TCAAAACCTT AATTTGGATT GTTAATTTAA TTTCCGCTTT
     51  GGCCGTCTTC GTGTTAGTAT TGCTCCAACA CGGCAAAGGC GCGGATGCCG
    101  GCGCGACTTT CGGA...
```

This corresponds to the amino acid sequence <SEQ ID 508; ORF 132>:

```
m132.pep  (partial)
      1  MEPFKTLIWI VNLISALAVF VLVLLQHGKG ADAGATFG...
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 132 shows 89.5% identity over a 38 aa overlap with a predicted ORF (ORF 132.ng) from *N. gonorrhoeae:*

```
m132/g132

                       10        20        30
m132.pep     MEPFKTLIWIVNLISALAVFVLVLLQHGKGADAGATFG
             || |||||||:|:||||||:|||||||||||||||||
g132         MEAFKTLIWIINIISALAVIVLVLLQHGKGADAGATFGSGSGSAQGVFGSAGNANFLSRS
                       10        20        30        40        50        60
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 509>:

```
a132.seq
      1  ATGGAAGCCT TCAAAACCCT AATTTGGATT GTTAATATAA TTTCCGCTTT
     51  GGCCGTCATC GTGTTAGTAT TGCTCCAACA CGGCAAAGGC GCGGATGCCG
    101  GCGCGACTTT CGGATCGGGA AGCGGCAGCG CGCAAGGCGT ATTCGGCTCT
    151  GCCGGCAACG CTAACTTCCT CAGCCGCTCG ACCGCCGTTG CAGCAACATT
    201  TTTCTTTGCA ACCTGCATGg GCTATGGTGT ATATTCACAC CCACACGACA
    251  AAACACGGTT TGGACTTCAG CAACGTACAA CAAACTCAGC AAGCACCCAA
    301  ACCCGTAAGC AATACCGAAC CTTCTGCCCC TGTTCCTCAG CAGCAGAAAT
    351  AACAGTTTTT CAAATGCCGA CATGGTGA
```

This corresponds to the amino acid sequence <SEQ ID 510; ORF 132.a>:

```
a132.pep
      1  MEAFKTLIWI VNIISALAVI VLVLLQHGKG ADAGATFGSG SGSAQGVFGS
     51  AGNANFLSRS TAVAATFFFA TCMGYGVYSH PHDKTRFGLQ QRTTNSASTQ
    101  TRKQYRTFCP CSSAAEITVF QMPTW*
```

m132/a132   92.1% identity in 38 aa overlap

```
201   TTCCGTGGCA TCAAGTGTGA TGCAGTTCGA TTACAAAGAC CACACCGTCA
251   ACCTCTTGGA CACGCCGGGA CACCAAGACT TCTCCGAAGA CACCTACCGC
301   GTTTTAACCG CCGTGGACAG CGCATTAATG GTCATCGACG CGGCAAAAGG
351   CGTGGAAGCG CAAACCATCA AGCTCTTAAA CGTCTGCCGC CTGCGCGATA
401   CACCGATTGT TACGTTTATG AACAAATACG ACCGCGAAGT GCGCGATTCC
451   CTGGAACTTT TGGACGAAGT GGAAAACATT TTAAAAATCC GCTGCGCGCC
501   CGTTACCTGG CCGATCGGTA TGGGCAAAAA CTTCAAGGGC GTGTACCACA
551   TCCTGAACGA TGAAATTTAT CTCTTTGAAG CTGGCGGCGA ACGCCTGCCG
601   CACGAGTTCG ACATCATCAA AGGCATCGAT AATCCTGAAT TGGAACAACG
651   CTTTCCGTTG GAAATCCAGC AGTTGCGCGA CGAAATCGAA TTGGTGCAGG
701   CGGCTTCCAA CGAGTTTAAT CTCGACGAAT TCCTCGCCGG CGAACTCACG
751   CCCGTATTCT TCGGCTCTGC GATTAACAAC TTCGGTATTC AGGAAATCCT
801   CAATTCATTG ATTGACTGGG CGCCCGCGCC GAAACCGCGC GACGCGACCG
851   TACGTATGGT CGAGCCGGAC GAGCCGAAGT TTTCCGGATT TATCTTCAAA
901   ATCCAAGCCA ATATGGACCC GAAACACCGC GACCGTATTG CCTTCTTGCG
951   CGTCTGCTCC GGCAAATTCG AGCGCGGCAT GAAGATGAAA CACCTGCGTA
1001  TCAACCGCGA AATCGCCGCC TCCAGCGTGG TTACCTTCAT GTCGCACGAC
1051  CGCGAGCTGG TTGAAGAAGC CTACGCCGGC GACATTATCG GCATCCCGAA
1101  CCACGGCAAC ATCCAAATCG GCGACAGCTT CTCCGAAGGC GAACAACTGG
1151  CGTTCACCGG CATCCCATTC TTCGCACCCG AACTGTTCCG CAGCGTACGC
1201  ATCAAAAACC CGCTGAAAAT CAAACAACTG CAAAAAGGCT TGCAACAGCT
1251  CGGCGAAGAA GGCGCGGTGC AGGTGTTCAA ACCGATGAGC GGCGCGGATT
1301  TGATTTTGGG CGCGGTCGGC GTGTTGCAGT TTGAAGTCGT TACCTCGCGC
1351  CTCGCCAACG AATACGGCGT AGAAGCCGTG TTCGACAGCG CATCCATCTG
1401  GTCGGCGCGC TGGGTATCGT GCGACGACAA GAAAAAACTG GCTGAATTTG
1451  AAAAAGCCAA CGCGGGCAAC CTCGCCATCG ACGCAGGCGG CAACCTCGCC
1501  TACCTCGCCC CCAACCGCGT GAATTTGGGA CTCACGCAAG AACGTTGGCC
1551  GGACATCGTG TTCCACGAAA CACGCGAACA TTCGGTCAAA CTGTAA
```

This corresponds to the amino acid sequence <SEQ ID 514; ORF 134>:

```
m134.pep
  1   MSQEILDQVR RRRTFAIISH PDAGKTTLTE KLLLFSGAIQ SAGTVKGKKT
 51   GKFATSDWME IEKQRGISVA SSVMQFDYKD HTVNLLDTPG HQDFSEDTYR
101   VLTAVDSALM VIDAAKGVEA QTIKLLNVCR LRDTPIVTFM NKYDREVRDS
151   LELLDEVENI LKIRCAPVTW PIGMGKNFKG VYHILNDEIY LFEAGGERLP
201   HEFDIIKGID NPELEQRFPL EIQQLRDEIE LVQAASNEFN LDEFLAGELT
251   PVFFGSAINN FGIQEILNSL IDWAPAPKPR DATVRMVEPD EPKFSGFIFK
301   IQANMDPKHR DRIAFLRVCS GKFERGMKMK HLRINREIAA SSVVTFMSHD
351   RELVEEAYAG DIIGIPNHGN IQIGDSFSEG EQLAFTGIPF FAPELFRSVR
401   IKNPLKIKQL QKGLQQLGEE GAVQVFKPMS GADLILGAVG VLQFEVVTSR
451   LANEYGVEAV FDSASIWSAR WVSCDDKKKL AEFEKANAGN LAIDAGGNLA
501   YLAPNRVNLG LTQERWPDIV FHETREHSVK L*
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

ORF 134 shows 98.7% identity over a 531 aa overlap with a predicted ORF (ORF 134.ng) from *N. gonorrhoeae:*

```
m134/g134

                    10        20        30        40        50        60
m134.pep    MSQEILDQVRRRRTFAIISHPDAGKTTLTEKLLLFSGAIQSAGTVKGKKTGKFATSDWME
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
g134        MSQEILDQVRRRRTFAIISHPDAGKTTLTEKLLLFSGAIQSAGTVKGKKTGKFATSDWMD
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m134.pep    IEKQRGISVASSVMQFDYKDHTVNLLDTPGHQDFSEDTYRVLTAVDSALMVIDAAKGVEA
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g134        IEKQRGISVASSVMQFDYKDHTVNLLDTPGHQDFSEDTYRVLTAVDSALMVIDAAKGVEA
                    70        80        90       100       110       120

                   130       140       150       160       170       180
```

```
 951  CGTCTGCTCC GGCAAATTCG AGCGCGGCAT GAAAATGAAA CACCTGCGTA
1001  TCAACCGCGA AATCGCCGCC TCCAGCGTGG TAACCTTCAT GTCCCACGAC
1051  CGCGAGCTGG TTGAAGAAGC CTACGCCGGC GACATTATCG GTATCCCAAA
1101  CCACGGCAAC ATCCAAATCG GCGACAGCTT CTCCGAAGGC GAACAACTGA
1151  CGTTTACCGG CATCCCATTC TTCGCGCCCG AACTGTTCCG CAGCGTTCGC
1201  ATCAAAAACC CGCTGAAAAT CAAGCAACTG CAAAAAGGTT TGCAACAGCT
1251  TGGCGAAGAA GGTGCGGTGC AGGTGTTCAA ACCAATGAGC GGCGCGGATT
1301  TGATTTTGGG CGCGGTCGGC GTGTTGCAGT TTGAAGTCGT TACCTCGCGC
1351  CTTGCCAACG AATACGGCGT GGAAGCCGTG TTCGACAACG CATCCATCTG
1401  GTCGGCGCGC TGGGTATCGT GCGACGACAA GAAAAAACTG GCGGAATTTG
1451  AAAAAGCCAA CGCGGGCAAC CTCGCCATCG ACGCGGGCGG CAACCTCGCC
1501  TACCTCGCCC CTAACCGCGT GAATCTGGGA CTCACGCAAG AACGCTGGCC
1551  GGACATCGTG TTCCACGAAA CGCGCGAGCA TTCGGTCAAA CTTTAA
```

This corresponds to the amino acid sequence <SEQ ID 516; ORF 134.a>:

```
a134.pep
  1  MSQEILDQVR RRRTFAIISH PDAGKTTLTE KLLLFSGAIQ SAGTVKGKKT
 51  GKFATSDWMD IEKQRGISVA SSVMQFDYKD HTVNLLDTPG HQDFSEDTYR
101  VLTAVDSALM VIDAAKGVEA QTIKLLNVCR LRNTPIVTFM NKYDREVRDS
151  LELLDEVENI LQIRCAPVTW PIGMGKNFKG VYHILNDEIY LFEAGGERLP
201  HEFDIIKGID NPELEQRFPL EIQQLRDEIE LVQAASNEFN LDEFLAGELT
251  PVFFGSAINN FGIQEILNSL IEWAPAPKPR DATVRMVEPD EPKFSGFIFK
301  IQANMDPKHR DRIAFLRVCS GKFERGMKMK HLRINREIAA SSVVTFMSHD
351  RELVEEAYAG DIIGIPNHGN IQIGDSFSEG EQLTFTGIPF FAPELFRSVR
401  IKNPLKIKQL QKGLQQLGEE GAVQVFKPMS GADLILGAVG VLQFEVVTSR
451  LANEYGVEAV FDNASIWSAR WVSCDDKKKL AEFEKANAGN LAIDAGGNLA
501  YLAPNRVNLG LTQERWPDIV FHETREHSVK L*
```

m134/a134 98.9% identity in 531 aa overlap

```
                    10        20        30        40        50        60
m134.pep    MSQEILDQVRRRRTFAIISHPDAGKTTLTEKLLLFSGAIQSAGTVKGKKTGKFATSDWME
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
a134        MSQEILDQVRRRRTFAIISHPDAGKTTLTEKLLLFSGAIQSAGTVKGKKTGKFATSDWMD
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m134.pep    IEKQRGISVASSVMQFDYKDHTVNLLDTPGHQDFSEDTYRVLTAVDSALMVIDAAKGVEA
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a134        IEKQRGISVASSVMQFDYKDHTVNLLDTPGHQDFSEDTYRVLTAVDSALMVIDAAKGVEA
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m134.pep    QTIKLLNVCRLRDTPIVTFMNKYDREVRDSLELLDEVENILKIRCAPVTWPIGMGKNFKG
            ||||||||||||:|||||||||||||||||||||||||||||:|||||||||||||||||
a134        QTIKLLNVCRLRNTPIVTFMNKYDREVRDSLELLDEVENILQIRCAPVTWPIGMGKNFKG
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m134.pep    VYHILNDEIYLFEAGGERLPHEFDIIKGIDNPELEQRFPLEIQQLRDEIELVQAASNEFN
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a134        VYHILNDEIYLFEAGGERLPHEFDIIKGIDNPELEQRFPLEIQQLRDEIELVQAASNEFN
                   190       200       210       220       230       240

                   250       260       270       280       290       300
m134.pep    LDEFLAGELTPVFFGSAINNFGIQEILNSLIDWAPAPKPRDATVRMVEPDEPKFSGFIFK
            |||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
a134        LDEFLAGELTPVFFGSAINNFGIQEILNSLIEWAPAPKPRDATVRMVEPDEPKFSGFIFK
                   250       260       270       280       290       300

                   310       320       330       340       350       360
m134.pep    IQANMDPKHRDRIAFLRVCSGKFERGMKMKHLRINREIAASSVVTFMSHDRELVEEAYAG
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a134        IQANMDPKHRDRIAFLRVCSGKFERGMKMKHLRINREIAASSVVTFMSHDRELVEEAYAG
                   310       320       330       340       350       360
```

```
 501  GACCGACATA GACGGTCTTT ACACGGGCAA CCCGAACAGC AATCCCGATG
 551  CCGTACGGCT GGACAAAATC GAACACATCA ACCATGAAAT CATCGAAATG
 601  GCGGGCGGCT CGGGTTCGGC AAACGGCACG GGCGGTATGC TGACCAAAAT
 651  CAAAGCGGCA ACCATCGCCG CCGAATCCGG CGTACCGGTG TATATCTGTT
 701  CCTCGCTCAA ACCCGATGCA CTTGCCGAAG CTGCCGAACA TCAGGCGGAC
 751  GGCTCGTTTT TCGTCCCCCG TGCCAAAGGT TTGCGGACGC AGAAGCAATG
 801  GCTGGCGTTC TATTCCGAAA GCCGGGGCAG CGTTTATGTG GACGAAGGTG
 851  CGGAACACGC TTTGTCCGAA CAGGGGAAAA GCCTGCTGAT GTCGGGCATT
 901  GCCGGAATCG AAGGGCATTT TTCCCGTATG GACACCGTAA CCGTGTACAG
 951  CAAGGCAACC AAACAGCCCC TGGGCAAAGG GCGCGTCCTG TTCGGCTCTG
1001  CCGCCGCCGA AGACCTGCTC AAATCGCGTA AGGCGAAAGG CGTGTTCATC
1051  CATCGGGACG ACTGGATTTC CATCACGCCC GAAATACGCC TGCTTCTGAC
1101  CGAATTTTAG
```

This corresponds to the amino acid sequence <SEQ ID 520; ORF 135>:

```
m135.pep
  1  MKYKRIVFKV GTSSITHSDG SLSRGKIQTI TCQLAALHHA GHELVLVSSG
 51  AVAAGFGALG FKKRPVKIAD KQASAAVGQG LLMEEYTANL SSDGIVSAQI
101  LLSRADFADK RRYQNAGGAL SVLLQRRAVP IINENDTVSV EELKIGDNDT
151  LSAQVAAMIQ ADLLVLLTDI DGLYTGNPNS NPDAVRLDKI EHINHEIIEM
201  AGGSGSANGT GGMLTKIKAA TIAAESGVPV YICSSLKPDA LAEAAEHQAD
251  GSFFVPRAKG LRTQKQWLAF YSESRGSVYV DEGAEHALSE QGKSLLMSGI
301  AGIEGHFSRM DTVTVYSKAT KQPLGKGRVL FGSAAAEDLL KSRKAKGVFI
351  HRDDWISITP EIRLLLTEF*
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

ORF 135 shows 97.6% identity over a 294 aa overlap with a predicted ORF (ORF 135.ng) from *N. gonorrhoeae*:

```
m135/g135

                    10        20        30        40        50        60
m135.pep    MKYKRIVFKVGTSSITHSDGSLSRGKIQTITCQLAALHHAGHELVLVSSGAVAAGFGALG
            ||||||||||||||||:||||||||||||||| ||||||||||||||||||||||||||||
g135        MKYKRIVFKVGTSSITRSDGSLSRGKIQTITRQLAALHHAGHELVLVSSGAVAAGFGALG
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m135.pep    FKKRPVKIADKQASAAVGQGLLMEEYTANLSSDGIVSAQILLSRADFADKRRYQNAGGAL
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g135        FKKRPVKIADKQASAAVGQGLLMEEYTANLSSDGIVSAQILLSRADFADKRRYQNAGGAL
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m135.pep    SVLLQRRAVPIINENDTVSVEELKIGDNDTLSAQVAAMIQADLLVLLTDIDGLYTGNPNS
            ||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
g135        SVLLQRRAIPIINENDTVSVEELKIGDNDTLSAQVAAMIQADLLVLLTDIDGLYTGNPNS
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m135.pep    NPDAVRLDKIEHINHEIIEMAGGSGSANGTGGMLTKIKAATIAAESGVPVYICSSLKPDA
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
g135        NPDAVRLDKIEHINHEIIEMAGGSGSANGTGGMLTKIKAATIAAESGVPVYICSSLKPDS
                   190       200       210       220       230       240

                   250       260       270       280       290       300
m135.pep    LAEAAEHQADGSFFVPRAKGLRTQKQWLAFYSESRGSVYVDEGAEHALSEQGKSLLMSGI
            |||||||||||||||||||||||||||||||||| ||||||:||||||||||:
g135        LAEAAEHQADGSFFVPRAKGLRTQKQWLAFYSESGGSVYVDESAEHALSEQGKACX
                   250       260       270       280       290

                   310       320       330       340       350       360
```

```
                    250       260       270       280       290       300
m135.pep   LAEAAEHQADGSFFVPRAKGLRTQKQWLAFYSESRGSVYVDEGAEHALSEQGKSLLMSGI
           |||||::|||||||||||||||||||||||||||||:|||||||||||||||||||||||
a135       LAEAADNQADGSFFVPRAKGLRTQKQWLAFYSESRGGVYVDEGAEHALSEQGKSLLMSGI
                    250       260       270       280       290       300

                    310       320       330       340       350       360
m135.pep   AGIEGHFSRMDTVTVYSKATKQPLGKGRVLFGSAAAEDLLKSRKAKGVFIHRDDWISITP
           ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
a135       AGIEGHFSRMDTVTVYSKATKQPLGKGRVLFGSAAAEDLLKLRKAKGVFIHRDDWISITP
                    310       320       330       340       350       360

                    370
m135.pep   EIRLLLTEFX
           ||||||||||
a135       EIRLLLTEFX
                    370
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 523>:

```
g136.seq
      1  ATGGAAATCC GGTTTCAGAC AGCATTTTTA CGTTTGGTTC AGatgaAAAC
     51  AAACGCTtca aTTCTtaccg caACACGCCT TGTATTTCCT GccgCTGCCG
    101  CACGGACAGG GATCGTTCCT GCCGGtTTTT TCCCCTTCCC TGCGGACGGT
    151  TTGCGGTTTG TTGATGACCG CCTGCCAGTA GCGGTAGATG TCtgccagcg
    201  cgTAAGGCag tTCGGAcgca agttccgcca gctcgccttc ggTGAATTGC
    251  AGgcggataa cgccgtttTC CTCTTCGTCg taaatgccgc ccactgccat
    301  cacgGGGTAA AACAGCTCTT CAAACGCTTC ATCATCGGCG GCTTCAAACC
    351  AATCGGTCGG CACAATGTCC AAACCGTAAA GATAGGCGTT GCACCAAGTG
    401  TAAAAATCGC TGCCGCCCTC GCCGTCGTCG TAGAGCCACA AATCGGGCAG
    451  CTTTTTATCC GACATCGCGG CGGTTGTTTC CATCGCCATT GCCAAAACCA
    501  GCCGTTCGAT TTCGGAACGT TCGGCGGCGG TAAATTGCGA TTCGTCGCCC
    551  AACACTTCGG GCAGCCAGTC GAGCGGTGCC AATTTGTCCG GCCCGCTCAA
    601  CAGCGCCGTC ATAAAACCTT GAACCTCGTC GCAACGCATC GTGTTGCCTT
    651  GTTCGCTTTT GGCATCCAAT AA
```

This corresponds to the amino acid sequence <SEQ ID 524; ORF 136.ng>:

```
g136.pep
      1  MEIRFQTAFL RLVQMKTNAS ILTATRLVFP AAAARTGIVP AGFFPFPADG
     51  LRFVDDRLPV AVDVCQRVRQ FGRKFRQLAF GELQADNAVF LFVVNAAHCH
    101  HGVKQLFKRF IIGGFKPIGR HNVQTVKIGV APSVKIAAAL AVVVEPQIGQ
    151  LFIRHRGGCF HRHCQNQPFD FGTFGGKLR FVAQHFGQPV ERCQFVRPAQ
    201  QRRHKTLNLV ATHRVALFAF GIQ*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 525>:

```
m136.seq
      1  ATGGAAACAA ACGCTTCAAT TCTTACCGCA ACACGCCTTG TATTTTCTGC
     51  CGCTGCCGCA CGGACAGGGA TCGTTCCTGC CTGTTTTTTC GCCTTCCCTG
    101  CGGACGGTTT GCGGTTTGTT GATGACTGCC TGCCAGTAGC GGTAGATATC
    151  CGCCAATGCA TAAGGCAACT CGGATTCCAG TTCCGCCAGC TCGCCTTCTG
    201  TGAATTGCAG ACGGATAGCG CCGTTTTCCT CTTCGTCGTA AATACCGCCC
    251  AATGCCATGA TGGGATAAAA CAACTCTTCA AACGCTTCAT CATCGACGGC
    301  TTCAAACCAA TCGGTCGGCA CAATATCCAA ACCGTAAAGA TAAGCATTGC
    351  ACCATGTGTA AAAATCGCTG CCGCCGTCTT CGTTTTCATA CAGCCACAAA
    401  TCGGGCAGTT TTTTATCCGA CATCGCGGCG GTTGTTTCCA TCGCCATTGC
    451  CAAAACCAGC CGTTCGATTT CGGAACGTTC GGCGGCGGTA AATTGCGATT
    501  CGTCGCCCAA CACTTCGGGC AGCCAGTCGA GCGGTGTCAA TTTGTCCGGC
    551  CCGCTCAACA GCGCCGTCAT AAAACCTTGA ACCTCGTCGC AACGCATCGT
    601  GTTGCCTTGT TCGCTTTTGG CATCCAACAA TTCGCTCAAC CGCCGTTTGG
    651  ATGCTTCGGT AAATTTTCGG GAATCCATCA TTTTCCTTTT CAAATGGGTT
    701  TTGCGCCCTA TTATCGCCGC AATGCCGTCT GA
```

This corresponds to the amino acid sequence <SEQ ID 526; ORF 136>:

```
  1  METNASILTA TRLVFSAAAA RTGIVPACFF AFPADGLRLV DDRLPVAVDI
 51  RQCIRQLGFQ FRQLAFCELQ TDSAVVLFVV NTAQCHDGIK QLFKRFIIDG
101  FKPIGRHNIQ TVKISIAPCV KIAAAVFVFI QPQIGQFFIR HRGGCFHRHC
151  QNQPFDFGTF GGGKLRFVAQ HFGQPVERCQ FVRPAQQRRH KTLNLVATHR
201  VALFAFGIQQ FAQPPFGCFG KFSGIHHFPF PMGFAPYYSG LNLNQDKATK
251  PQTVQIVRQG EATPYWFKFN PLYRRNAV*
```

m136/a136 98.3% identity in 238 aa overlap

```
                    10        20        30        40        50        60
m136.pep    METNASILTATRLVFSAAAARTGIVPACFFAFPADGLRFVDDCLPVAVDIRQCIRQLGFQ
            ||||||||||||||||||||||||||||||||||||||||:||| |||||||||||||||
a136        METNASILTATRLVFSAAAARTGIVPACFFAFPADGLRLVDDRLPVAVDIRQCIRQLGFQ
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m136.pep    FRQLAFCELQTDSAVFLFVVNTAQCHDGIKQLFKRFIIDGFKPIGRHNIQTVKISIAPCV
            ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
a136        FRQLAFCELQTDSAVVLFVVNTAQCHDGIKQLFKRFIIDGFKPIGRHNIQTVKISIAPCV
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m136.pep    KIAAAVFVFIQPQIGQFFIRHRGGCFHRHCQNQPFDFGTFGGGKLRFVAQHFGQPVERCQ
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a136        KIAAAVFVFIQPQIGQFFIRHRGGCFHRHCQNQPFDFGTFGGGKLRFVAQHFGQPVERCQ
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m136.pep    FVRPAQQRRHKTLNLVATHRVALFAFGIQQFAQPPFGCFGKFSGIHHFPFQMGFAPYYRR
            |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
a136        FVRPAQQRRHKTLNLVATHRVALFAFGIQQFAQPPFGCFGKFSGIHHFPFPMGFAPYYSG
                   190       200       210       220       230       240


m136.pep    NAVX

a136        LNLNQDKATKPQTVQIVRQGEATPYWFKFNPLYRRNAVX
                   250       260       270
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 529>:

```
g137.seq
  1  ATGATTATCC ATCACcaaTT CGATCCCGTC CTCATCAGTA TCGGCCCGCT
 51  TGCCGTCCGC TGGTATGCCT TAAGCTACAT CCTCGGATTT ATTCTTTTTA
101  CCTTTCTCGG CAGAAGGCGC ATCGCGCAAG GCTTGTCCGT TTTTACCAAA
151  GAATCGCTCG ACGACTTCCT GACATGGGGC ATTTTGGGCG TGATTTTGGG
201  CGGACGCTTG GGCTATGTCC TGTTTTACAA ATTCTCCGAC TACCTCGCCC
251  ATCCGCTTGA TATTTTCAAG GTATGGGAAG GCGGAATGTC GTTCCACGGC
301  GGCTTTTTGG GTGTAGTTAT TGCCATATGG TTGTTCAGCC GCAAGCACGG
351  CATCGGCTTC CTCAAACTGA TGGACACGGT CGCGCCGCTC GTTCCGCTGG
401  GTCTCGCTTC GGGACGTATC GGCAACTTTA TCAACGGCGA ACTTTGGGGA
451  CGCATTACCG ACATTAACGC ATTTTGGGCA ATGGGCTTCC CGCAAGCGCA
501  TTACGAAGAT GCCGAAGCCG CCGCGCACAA TCCGCTTTGG GCAGAATGGC
551  TGCAACAATA CGGTATGCTG CCGCGTCATC CCTCGCAGCT TTATCAGTTT
601  GCCCTTGAAG GCATCTGCCT GTTCGCCGTC GTTTGGCTGT TTTCCAAAAA
651  ACCGCGCCCG ACCGGGCAGA CTGCCGCGCT TTTTCTCGGC GGCTACGGCG
701  TGTTCCGCTT TATTGCCGAA TTTGCGCGCC AACCCGACGA CTATCTCGGG
751  CTGCTGACCT TGGGGCTGTC GATGGGGCAA TGGTTGAGCG TCCCGATGAT
801  TGTTTTGGGT ATCGTCGGCT TTGTCCGGTT CGGCATGAAA AAACAGCACT
851  GA
```

This corresponds to the amino acid sequence <SEQ ID 530; ORF 137.ng>:

```
g137.pep
```

```
              ||||||||||||||||||||||||||||:|:|||||| | |||:|:|||||||:||||||
g137          AEWLQQYGMLPRHPSQLYQFALEGICLFAVVWLFSKKPRPTGQTAALFLGGYGVFRFIAE
                      190       200       210       220       230       240

                      250       260       270       280
m137.pep      FARQPDDYLGLLTLGLSMGQWLSVPMIVLGIVGFVRFGMKKQHX
              |||||||||||||||||||||||||||||||||||||||||||
g137          FARQPDDYLGLLTLGLSMGQWLSVPMIVLGIVGFVRFGMKKQHX
                      250       260       270       280
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 533>:

```
a137.seq
     1  ATGATTACCC ATCCCCAATT CGACCCCGTC CTTATCAGTA TCGGCCCGCT
    51  TGCCGTCCGC TGGTATGCCC TAAGCTACAT CCTCGGATTT ATTCTTTTTA
   101  CCTTTCTCGG CAGAAGGCGC ATCGCGCAAG GCTTGTCCGT TTTTACCAAA
   151  GAATCGCTCG ACGACTTCCT GACATGGGGC ATTTTGGGCG TAATTTTGGG
   201  CGGGCGTTTG GGTTACGTCC TGTTTTACAA GTTTTCCGAC TACCTCGCCC
   251  ATCCGCTTGA TATTTTCAAG GTATGGGAAG GCGGAATGTC GTTCCACGGC
   301  GGCTTTTTGG GTGTAGTTAT TGCCATATGG TTGTTCGGTC GCAAACACGG
   351  CATCGGCTTC CTCAAACTGA TGGACACGGT CGCACCGCTC GTTCCACTGG
   401  GTCTCGCTTC GGGACGTATC GGCAACTTCA TCAACGGCGA ACTTTGGGGA
   451  CGCGTTACCG ACATCAACGC ATTTTGGGCA ATGGGCTTCC CGCAGGCGCG
   501  TTACGAAGAC CTCGAAGCCG CCGCGCACAA TCCGCTTTGG GCAGAATGGC
   551  TGCAACAATA CGGTATGCTG CCGCGTCATC CCTCGCAGCT TTATCAGTTT
   601  GCACTTGAAG GCATCTGCCT GTTCGCCGTC GTTTGGCTGT TCTCTAAAAA
   651  ACAGCGGCCG ACCGGACAAG TCGCCTCACT CTTCCTCGGC GGCTACGGCA
   701  TATTCCGCTT CATTGCCGAA TTTGCACGCC AACCCGACGA CTATCTCGGG
   751  CTGCTGACCT TGGGGCTGTC GATGGGGCAA TGGTTGAGCG TCCCGATGAT
   801  TGTTTTGGGT ATCGTCGGCT TTGTCCGGTT CGGCATGAAA AAACAGCACT
   851  GA
```

This corresponds to the amino acid sequence <SEQ ID 534; ORF 137.a>:

```
a137.pep
     1  MITHPQFDPV LISIGPLAVR WYALSYILGF ILFTFLGRRR IAQGLSVFTK
    51  ESLDDFLTWG ILGVILGGRL GYVLFYKFSD YLAHPLDIFK VWEGGMSFHG
   101  GFLGVVIAIW LFGRKHGIGF LKLMDTVAPL VPLGLASGRI GNFINGELWG
   151  RVTDINAFWA MGFPQARYED LEAAAHNPLW AEWLQQYGML PRHPSQLYQF
   201  ALEGICLFAV VWLFSKKQRP TGQVASLFLG GYGIFRFIAE FARQPDDYLG
   251  LLTLGLSMGQ WLSVPMIVLG IVGFVRFGMK KQH*
```

m137/a137   98.2% identity in 283 aa overlap

```
                       10        20        30        40        50        60
m137.pep      MITHPQFDPVLISIGPLAVRWYALSYILGFILFTFLGRRRIAQGLSVFTKESLDDFLTWG
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a137          MITHPQFDPVLISIGPLAVRWYALSYILGFILFTFLGRRRIAQGLSVFTKESLDDFLTWG
                       10        20        30        40        50        60

                       70        80        90       100       110       120
m137.pep      ILGVILGGRLGYVLFYKFSDYLAHPLDIFKVWEGGMSFHGGFLGVVIAIRLFGRKHGIGF
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
a137          ILGVILGGRLGYVLFYKFSDYLAHPLDIFKVWEGGMSFHGGFLGVVIAIWLFGRKHGIGF
                       70        80        90       100       110       120

                      130       140       150       160       170       180
m137.pep      LKLMDTVAPLVPLGLASGRIGNFINGELWGRVTDINAFWAMGFPQARYEDAEAAAHNPLW
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
a137          LKLMDTVAPLVPLGLASGRIGNFINGELWGRVTDINAFWAMGFPQARYEDLEAAAHNPLW
                      130       140       150       160       170       180

                      190       200       210       220       230       240
m137.pep      AEWLQQYGMLPRHPSQLYQFALEGICLFTVIWLFSKKQRSTGQVASLFLGGYGIFRFIAE
              ||||||||||||||||||||||||||||:|:|||||||| ||||||||||||||||||||
a137          AEWLQQYGMLPRHPSQLYQFALEGICLFAVVWLFSKKQRPTGQVASLFLGGYGIFRFIAE
```

```
 51  RDVVLLKLVG IHPVIVHGGG PQINAMLEKV GKKGEFVQGM RVTDKEAMDI
101  VEMVLGGHVN KEIVSMINTY GGHAVGVSGR DDHFIKAKKL LIDTPEQNGV
151  DIGQVGTVES IDTGLVKGLI ERGCIPVVAP VGVGEKGEAF NINADLVAGK
201  LAEELNAEKL LMMTNIAGVM DKTGNLLTKL TPKRIDELIA DGTLYGGMLP
251  KIASAVEAAV NGVKATHIID GRLPNALLLE IFTDAGIGSM ILGGGEDA*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 138 shows 98.0% identity over a 298 aa overlap with a predicted ORF (ORF 138.ng) from *N. gonorrhoeae:*

```
m138/g138

                  10        20        30        40        50        60
m138.pep  MESENIISAADKARILAEALPYIRRFSGSVAVIKYGGNAMTEPALKEGFARDVVLLKLVG
          || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g138      MEFENIISAADKARILAEALPYIRRFSGSVAVIKYGGNAMTEPALKEGFARDVVLLKLVG
                  10        20        30        40        50        60

                  70        80        90       100       110       120
m138.pep  IHPVIVHGGGPQINAMLEKVGKKGEFVQGMRVTDKEAMDIVEMVLGGHVNKEIVSMINTY
          ||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
g138      IHPVIVHGGGPQINAMLEKVGKKGEFVQGMRVTDKETMDIVEMVLGGHVNKEIVSMINTY
                  70        80        90       100       110       120

                 130       140       150       160       170       180
m138.pep  GGHAVGVSGRDDHFIKAKKLLIDTPEQNGVDIGQVGTVESIDTGLVKGLIERGCIPVVAP
          ||||||||||||||||||||:||||||:||||||||||||||||||||||||||||||||
g138      GGHAVGVSGRDDHFIKAKKLLVDTPEQNSVDIGQVGTVESIDTGLVKGLIERGCIPVVAP
                 130       140       150       160       170       180

                 190       200       210       220       230       240
m138.pep  VGVGEKGEAFNINADLVAGKLAEELNAEKLLMMTNIAGVMDKTGNLLTKLTPKRIDELIA
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
g138      VGVGEKGEAFNINADLVAGKLAEELNAEKLLMMTNIAGVMDKTGNLLTKLTPKRIDGLIA
                 190       200       210       220       230       240

                 250       260       270       280       290      299
m138.pep  DGTLYGGMLPKIASAVEAAVNGVKATHIIDGRLPNALLLEIFTDAGIGSMILGGGEDAX
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
g138      DGTLYGGMLPKIASAVEAAVNGVKATHIIDGRLPNALLLEIFTDAGIGSMILGRGEDAX
                 250       260       270       280       290
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 539>:

```
a138.seq
   1  ATGGAGTCTG AAAACATTAT TTCCGCCGCC GACAAGGCGC GTATCCTTGC
  51  CGAAGCGCTG CCTTACATCC GCCGGTTTTC CGGTTCGGTC GCCGTCATCA
 101  AATACGGCGG CAACGCGATG ACCGAACCTG CCTTGAAAGA AGGGTTTGCC
 151  CGCGATGTCG TGCTGCTGAA GCTGGTCGGC ATTCATCCCG TCATCGTTCA
 201  CGGCGGCGGG CCGCAGATCA ATGCGATGCT TGAAAAAGTC GGCAAAAAGG
 251  GTGAGTTTGT CCAAGGAATG CGCGTTACCG ACAAAGAGGC GATGGATATT
 301  GTCGAAATGG TGTTGGGCGG GCATGTCAAT AAAGAAATCG TGTCGATGAT
 351  TAACACATAT GGCGGACACG CGGTCGGCGT AAGCGGACGC GACGACCATT
 401  TCATTAAGGC GAAGAAACTT TTGATCGATA CGCCCGAACA GAATGGCGTG
 451  GACATCGGAC AGGTCGGTAC GGTGGAAAGC ATCGATACCG GTTTGGTTAA
 501  AGGGCTGATA GAACGTGGCT GCATTCCCGT CGTCGCCCCC GTCGGCGTAG
 551  GTGAAAAAGG CGAAGCGTTC AACATCAACG CCGATTTGGT AGCAGGCAAA
 601  TTGGCGGAAG AATTGAACGC CGAAAAACTC TTGATGATGA CGAATATCGC
 651  CGGTGTGATG GACAAAACGG GCAATCTGCT GACCAAACTC ACGCCGAAAC
 701  GGATTGATGA ACTGATTGCC GACGGCACGC TGTATGGCGG TATGCTGCCG
 751  AAAATCGCTT CTGCGGTCGA AGCCGCCGTC AACGGCGTGA AAGCCACGCA
 801  TATCATCGAC GGCAGGGTGC CCAACGCGCT TTTGCTGGAA ATCTTACCG
 851  ATGCCGGTAT CGGTTCGATG ATTTTGGGCG GTGGGGAAGA TGCCTGA
```

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 543>:

```
m139.seq
      1  ATGCGAACGA CCCCAACCTT CCCTACAAAA ACTTTCAAAC CGACTGCCAT
     51  GGCGTTAGCT GTTGCAACAA CACTTTCTGC CTGCTTAGGC GGCGGCGGAG
    101  GCGGCACTTC TGCGCCCGAC TTCAATGCAG GCGGTACCGG TATCGGCAGC
    151  AACAGCAGAG CAACAACAGC GAAATCAGCA GCAGTATCTT ACGCCGGTAT
    201  CAAGAACGAA ATGTGCAAAG ACAGAAGCAT GCTCTGTGCC GGTCGGGATG
    251  ACGTTGCGGT TACAGACAGG GATGCCAAAA TCAATGCCCC CCCCCGAATC
    301  TGCATACCGG AGACTTTCCA AACCCAAATG ACGCATtACA AGAATTTGAT
    351  CAACCTCAAA CCTGCAATTG AAGCAGGCTA TACAGGACGC GGGGTAGAGG
    401  TAGGTATCGT CGACACAGGC GAATCCGTCG GCAGCATATC CTTTCCCGAA
    451  CTGTATGGCA GAAAAGAACA CGGCTATAAC GAAAATTACG AAAAACTATA
    501  CGGCGTATAT GCGGAAGGAA GCGCCTGA
```

This corresponds to the amino acid sequence <SEQ ID 544; ORF 138>:

```
m139.pep
      1  MRTTPTFPTK TFKPTAMALA VATTLSACLG GGGGGTSAPD FNAGGTGIGS
     51  NSRATTAKSA AVSYAGIKNE MCKDRSMLCA GRDDVAVTDR DAKINAPPRI
    101  CIPETFQTQM THYKNLINLK PAIEAGYTGR GVEVGIVDTG ESVGSISFPE
    151  LYGRKEHGYN ENYEKLYGVY AEGSA*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 138 shows 92.2% identity over a 179 aa overlap with a predicted ORF (ORF 138.ng) from *N. gonorrhoeae:*

```
m139/g139

                           10        20        30        40        50        60
m139.pep          MRTTPTFPTKTFKPTAMALAVATTLSACLGGGGGGTSAPDFNAGGTGIGSNSRATTAKSA
                  |||| |||||||||:|||||||||||||||||||||||||||||||||||||||| |:||
g139              MRTTSTFPTKTFKPAAMALAVATTLSACLGGGGGGTSAPDFNAGGTGIGSNSRATIAESA
                           10        20        30        40        50        60

                           70        80        90       100       110       120
m139.pep          AVSYAGIKNEMCKDRSMLCAGRDDVAVTDRDAKINAPPRICIPETFQTQMTHYKNLINLK
                  ||||||||||||||||||||||||||||||||||:|| |||||||||||||: ||:||||
g139              AVSYAGIKNEMCKDRSMLCAGRDDVAVTDRDAKIKAP-RICIPETFQTQMTNIKNMINLK
                           70        80        90        100       110

                          130       140       150       160          170
m139.pep          PAIEAGYTGRGVEVGIVDTGESVGSISFPELYGRKEHGYNENY----EKLYGVYAEGSAX
                  |||||||||||||||||||||||||||||||||||||||||||    :||||||||||||
g139              PAIEAGYTGRGVEVGIVDTGESVGSISFPELYGRKEHGYNENYKNKLQKLYGVYAEGSAX
                      120       130       140       150       160       170
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 545>:

```
a139.seq
      1  ATGCGAACGA CCCCAACCTT CCCTACAAAA ACTTTCAAAC CGGCTGCCAT
     51  GGCGTTAGCT GTTGCAACAA CACTTTCTGC CTGCTTAGGC GGCGGCGGAG
    101  GCGGCACTTC TGCGCCCGAC TTCAATGCAG GCGGCACCGG TATCGGCAGC
    151  AACAGCAGGG CAACAACAGC GAAATCAGCA GCAATATCTT ACGCCGGTAT
    201  CAAGAACGAA ATGTGCAAAG ACAGAAGCAT GCTCTGTGCC GGTCGGGATG
    251  ACGTTGCGGT TACAGACAGG GATGCCAAAA TCAATGCCCC CCCCCGAATC
    301  TGCATACCGG AGACTTTACA AACCCAAATG ACGCAT.ACA AGAATTTGAT
    351  CAACCTCAAA CCTGCAATTG AAGCAGGCTA TACAGGACGC GGGGTAGAGG
    401  TAGGTATCGT CGACACAGGC GAATCCGTCG GCAGCATATC CTTTCCCGAA
    451  CTGTATGGCA GAAAAGAACA CGGCTATAAC GAAAATTAC. AAAAACTATA
    501  CGGCGTATAT GCGGAAGGAA GCGCCTGA
```

This corresponds to the amino acid sequence <SEQ ID 546; ORF 139.a>:

```
a139.pep
      1  MRTTPTFPTK TFKPAAMALA VATTLSACLG GGGGGTSAPD FNAGGTGIGS
```

```
251  DSISLFAGIR HDVGDIGYLK GLFSYGRYKN SISRSTGADE YAEGSVNGTL
301  MQLGALGGVN VPFAATGDLT VEGGLRHDLL KQDAFAEKGS ALGWSGNSLT
351  EGTLVGLAGL KLSQPLSDKA VLSATAGVER DLNGRDYAVT GGFTGAAAAT
401  GKTGARNMPH TRRVAGLGVD VEFGNGWNGL ARYSYTGSKQ YGNHSGQIGV
451  GYRF*
```

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 549>:

```
m140.seq
   1  ATGTCGGCAC GCGGCAAGGG GGCAGGCTAT CTCAACAGTA CCGGACGACG
  51  TGTTCCCTTC CTGAGTGCCG CCAAAATCGG GCAGGATTAT TCTTTCTTCA
 101  CAAACATCGA AACCGACGGC GGCCTGCTGG CTTCCCTCGA CAGCGTCGAA
 151  AAAACAGCGG GCAGTGAAGG CGACACGCTG TCCTATTATG TCCGTCGCGG
 201  CAATGCGGCA CGGACTGCTT CGGCAGCGGC ACATTCCGCG CCCGCCGGTC
 251  TGAAACACGC CGTAGAACAG GGCGGCAGCA ATCTGGAAAA CCTGATGGTC
 301  GAACTGGATG CCTCCGAATC ATCCGCAACA CCCGAGACGG TTGAAACTGC
 351  GGCAGCCGAC CGCACAGATA TGCCGGGCAT CCGCCCCTAC GGCGCAACTT
 401  TCCGCGCAGC GGCAGCCGTA CAGCATGCGA ATGCCGCCGA CGGTGTACGC
 451  ATCTTCAACA GTCTCGCCGC TACCGTCTAT GCCGACAGTA CCGCCGCCCA
 501  TGCCGATATG CAGGGACGCC GCCTGAAAGC CGTATCGGAC GGGTTGGACC
 551  ACAACGGCAC GGGTCTGCGC GTCATCGCGC AAACCCAACA GGACGGTGGA
 601  ACGTGGGAAC AGGGCGGTGT TGAAGGCAAA ATGCGCGGCA GTACCCAAAC
 651  CGTCGGCATT GCCGCGAAAA CCGGCGAAAA TACGACAGCA GCCGCCACAC
 701  TGGGCATGGG ACGCAGCACA TGGAGCGAAA ACAGTGCAAA TGCAAAAACC
 751  GACAGCATTA GTCTGTTTGC AGGCATACGG CACGATGCGG GCGATATCGG
 801  CTATCTCAAA GGCCTGTTCT CCTACGGACG CTACAAAAAC AGCATCAGCC
 851  GCAGCACCGG TGCGGACGAA CATGCGGAAG GCAGCGTCAA CGGCACGCTG
 901  ATGCAGCTGG GCGCACTGGG CGGTGTCAAC GTTCCGTTTG CCGCAACGGG
 951  AGATTTGACG GTCGAAGGCG GTCTGCGCTA CGACCTGCTC AAACAGGATG
1001  CATTCGCCGA AAAAGGCAGT GCTTTGGGCT GGAGCGGCAA CAGCCTCACT
1051  GAAGGCACGC TGGTCGGACT CGCGGGTCTG AAGCTGTCGC AACCCTTGAG
1101  CGATAAAGCC GTCCTGTTTG CAACGGCGGG CGTGGAACGC GACCTGAACG
1151  GACGCGACTA CACGGTAACG GGCGGCTTTA CCGGCGCGAC TGCAGCAACC
1201  GGCAAGACGG GGGCACGCAA TATGCCGCAC ACCCGTCTGG TTGCCGGCCT
1251  GGGCGCGGAT GTCGAATTCG GCAACGGCTG GAACGGCTTG GCACGTTACA
1301  GCTACGCCGG TTCCAAACAG TACGGCAACC ACAGCGGACG AGTCGGCGTA
1351  GGCTACCGGT TCTGA
```

This corresponds to the amino acid sequence <SEQ ID 550; ORF 140>:

```
m140.pep
   1  MSARGKGAGY LNSTGRRVPF LSAAKIGQDY SFFTNIETDG GLLASLDSVE
  51  KTAGSEGDTL SYYVRRGNAA RTASAAAHSA PAGLKHAVEQ GGSNLENLMV
 101  ELDASESSAT PETVETAAAD RTDMPGIRPY GATFRAAAAV QHANAADGVR
 151  IFNSLAATVY ADSTAAHADM QGRRLKAVSD GLDHNGTGLR VIAQTQQDGG
 201  TWEQGGVEGK MRGSTQTVGI AAKTGENTTA AATLGMGRST WSENSANAKT
 251  DSISLFAGIR HDAGDIGYLK GLFSYGRYKN SISRSTGADE HAEGSVNGTL
 301  MQLGALGGVN VPFAATGDLT VEGGLRYDLL KQDAFAEKGS ALGWSGNSLT
 351  EGTLVGLAGL KLSQPLSDKA VLFATAGVER DLNGRDYTVT GGFTGATAAT
 401  GKTGARNMPH TRLVAGLGAD VEFGNGWNGL ARYSYAGSKQ YGNHSGRVGV
 451  GYRF*
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N.gonorrhoeae*</u>

ORF 140 shows 94.5% identity over a 454 aa overlap with a predicted ORF (ORF 140.ng) from *N. gonorrhoeae:*

```
m140/g140

                   10        20        30        40        50        60
m140.pep   MSARGKGAGYLNSTGRRVPFLSAAKIGQDYSFFTNIETDGGLLASLDSVEKTAGSEGDTL
           ||||||||||||||||:|||||||||||||||| ||:|||||||||||||||||||||| 
g140       MSARGKGAGYLNSTGRHVPFLSAAKIGQDYSFFKNIKTDGGLLASLDSVEKTAGSEGDTP
                   10        20        30        40        50        60

                   70        80        90       100       110       120
```

```
 951  AGATTTGACG GTCGAAGGCG GTCTGCGCTA CGACCTGCTC AAACAGGATG
1001  CATTCGCCGA AAAAGGCAGT GCTTTGGGCT GGAGCGGCAA CAGCATCACT
1051  GAAGGCACAC TGGTCGGACT CGCGGGTCTG AAGCTGTCGC AACCCTTGAG
1101  CGATAAAGCC GTCCTGTTTG CAACGGCGGG CGTGGAACGC GACCTGAACG
1151  GACGCGACTA CACGGTAACG GGCGGCTTTA CCGGCGCGAC TGCAGCAACC
1201  GGCAAGACGG GGGCACGCAA TATGCCGCAC ACCCGCCTGG TTGCCGGTCT
1251  GGGCGCGGAT GTCGAATTCG GCAACGGCTG GAACGGCTTG GCACGTTACA
1301  GCTACGCCGG TTCCAAACAG TACGGCAACC ACAGCGGACG AGTCGGCGTA
1351  GGCTACCGGT TCTGA
```

This corresponds to the amino acid sequence <SEQ ID 552; ORF 140.a>:

```
a140.pep
  1  MSAGGKGAGY LNRTGQRVPF LSAAKIGRDY SFFTNIETDG GLLASLDSVE
 51  KTAGSEGDTL SYYVRRGNAA RTASAAAHSA PAGLKHAVEQ GGSNLENLMV
101  ELDASESSAT PETVETAAAD RTDMPGIRPY GATFRAAAAV QHANAADGVR
151  IFNNLAATVY ADSTAAHADM QGRRLKAVSD GLDHNATGLR VIAQTQQDGG
201  TWEQGGVEGK MRGSTQTVGI AAKTGENTTA AATLGMGHST WSENSANAKT
251  DSISLFAGIR HDAGDIGYLK GLFSYGRYKN SISRSTGADE HAEGSVNGTL
301  MQLGALGGVN VPFAATGDLT VEGGLRYDLL KQDAFAEKGS ALGWSGNSIT
351  EGTLVGLAGL KLSQPLSDKA VLFATAGVER DLNGRDYTVT GGFTGATAAT
401  GKTGARNMPH TRLVAGLGAD VEFGNGWNGL ARYSYAGSKQ YGNHSGRVGV
451  GYRF*
```

m140/a140 98.2% identity in 454 aa overlap

```
                 10        20        30        40        50        60
m140.pep   MSARGKGAGYLNSTGRRVPFLSAAKIGQDYSFFTNIETDGGLLASLDSVEKTAGSEGDTL
           ||| |||||||| ||:||||||||||:|||||||||||||||||||||||||||||||||
a140       MSAGGKGAGYLNRTGQRVPFLSAAKIGRDYSFFTNIETDGGLLASLDSVEKTAGSEGDTL
                 10        20        30        40        50        60

                 70        80        90       100       110       120
m140.pep   SYYVRRGNAARTASAAAHSAPAGLKHAVEQGGSNLENLMVELDASESSATPETVETAAAD
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a140       SYYVRRGNAARTASAAAHSAPAGLKHAVEQGGSNLENLMVELDASESSATPETVETAAAD
                 70        80        90       100       110       120

                130       140       150       160       170       180
m140.pep   RTDMPGIRPYGATFRAAAAVQHANAADGVRIFNSLAATVYADSTAAHADMQGRRLKAVSD
           |||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
a140       RTDMPGIRPYGATFRAAAAVQHANAADGVRIFNNLAATVYADSTAAHADMQGRRLKAVSD
                130       140       150       160       170       180

                190       200       210       220       230       240
m140.pep   GLDHNGTGLRVIAQTQQDGGTWEQGGVEGKMRGSTQTVGIAAKTGENTTAAATLGMGRST
           |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||:||
a140       GLDHNATGLRVIAQTQQDGGTWEQGGVEGKMRGSTQTVGIAAKTGENTTAAATLGMGHST
                190       200       210       220       230       240

                250       260       270       280       290       300
m140.pep   WSENSANAKTDSISLFAGIRHDAGDIGYLKGLFSYGRYKNSISRSTGADEHAEGSVNGTL
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a140       WSENSANAKTDSISLFAGIRHDAGDIGYLKGLFSYGRYKNSISRSTGADEHAEGSVNGTL
                250       260       270       280       290       300

                310       320       330       340       350       360
m140.pep   MQLGALGGVNVPFAATGDLTVEGGLRYDLLKQDAFAEKGSALGWSGNSLTEGTLVGLAGL
           ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
a140       MQLGALGGVNVPFAATGDLTVEGGLRYDLLKQDAFAEKGSALGWSGNSITEGTLVGLAGL
                310       320       330       340       350       360

                370       380       390       400       410       420
m140.pep   KLSQPLSDKAVLFATAGVERDLNGRDYTVTGGFTGATAATGKTGARNMPHTRLVAGLGAD
           |||||||||||'||||||||||||||||||||||||||||||||||||||||||||||||
a140       KLSQPLSDKAVLFATAGVERDLNGRDYTVTGGFTGATAATGKTGARNMPHTRLVAGLGAD
```

```
m141.seq
      1  ATGAGCTTCA AAACCGATGC CGAAATCGCC CAATCCTCCA CCATGCGCCC
     51  GATTGGCGAA ATTGCCGCCA AGCTTGGTCT GAATGCCGAC AACATTGAGC
    101  CTTACGGTCA TTACAAGGCG AAAATCAATC CTGCCGAAGC GTTCAAACTG
    151  CCGCAAAAAC AGGGCAGGCT GATTTTGGTT ACCGCCATCA ACCCGACTCC
    201  GGCGGGCGAA GGCAAAACCA CCGTAACCAT CGGTTTGGCG GACGCGTTGC
    251  GCCACATCGG CAAAGATGCC GTGATTGCCC TGCGCGAACC TTCTCTGGGG
    301  CCGGTGTTCG GCGTGAAAGG CGGCGCGGCA GGCGGCGGCT ATGCCCAAGT
    351  TTTGCCGATG GAAGACATCA ACCTGCACTT CACCGGAGAT TTTCACGCCA
    401  TCGGTGCGGC AAATAATCTG CTTGCCGCGA TGCTCGACAA CCATATCTAC
    451  CAAGGCAACG AGTTGAACAT CGACCCCAAA CGCGTGCTGT GGCGGCGCGT
    501  GGTCGATATG AACGACCGCC AGTTGCGCAA CATCATCGAC GGCATGGGTA
    551  AACCCGTTGA CGGCGTGATG CGTCCTGACG GTTTCGATAT TACCGTTGCT
    601  TCCGAAGTGA TGGCGGTATT CTGTCTTGCC AAAGACATCA GCGATTTGAA
    651  AGAGCGTTTG GGCAACATCC TTGTCGCCTA CGCCAAAGAC GGCAGCCCCG
    701  TTTACGCCAA AGATTTGAAA GCGAATGGCG CGATGGCGGC ATTGCTTAAA
    751  GATGCGATTA AGCCCAACTT GGTGCAAACC ATCGAAGGCA CGCCCGCCTT
    801  CGTACACGGC GGCCCGTTCG CCAACATCGC CCACGGCTGC AACTCCGTAA
    851  CCGCAACCCG TCTGGCGAAA CACCTTGCCG ATTACGCCGT AACCGAAGCA
    901  GGCTTCGGCG CGGACTTGGG CGCGGAAAAA TTCTGCGACA TCAAATGCCG
    951  CCTTGCCGGT TTGAAACCTG ATGCGGCTGT TGTCGTGGCG ACTGTCCGCG
   1001  CGTTGAAATA TAACGGCGGC GTGGAACGCG CCAACCTCGG CGAAGAAAAT
   1051  TTAGACGCTT TGGAAAAAGG TTTGCCCAAC CTGCTGAAAC ACATTTCCAA
   1101  CCTGAAAAAC GTATTCGGAC TGCCCGTCGT CGTTGCGCTC AACCGCTTCG
   1151  TGTCCGACGC CGATGCCGAG TTGGCGATGA TTGAAAAAGC CTGTGCCGAA
   1201  CACGGCGTTG AAGTTTCCCT GACCGAAGTG TGGGGCAAAG GTGGTGCGGG
   1251  CGGCGCGGAT TTGGCGCGCA AAGTCGTCAA CGCCATTGAA AGTCAAACCA
   1301  ATAACTTCGG TTTCGCCTAC GATGTCGAGT TGGGCATCAA AGACAAAATC
   1351  CGTGCGATTG CCCAAAAAGT GTACGGCGCG GAAGATGTTG ATTTCAGCGC
   1401  GGAAGCGTCT GCCGAAATCG CTTCACTGGA AAAACTGGGC TTGGACAAAA
   1451  TGCCGATCTG CATGGCGAAA ACCCAATACT CTTTGAGCGA CAACGCCAAA
   1501  CTGTTGGGCT GCCCCGAAGA CTTCCGCATC GCCGTGCGCG GCATCACCGT
   1551  TTCCGCAGGC GCAGGTTTCA TCGTCGCCCT GTGCGGCAAC ATGATGAAAA
   1601  TGCCCGGCCT GCCCAAAGTT CCGGCTGCCG AGAAAATCGA TGTGGACGCA
   1651  GAAGGCGTGA TTCACGGCTT GTTCTGA
```

This corresponds to the amino acid sequence <SEQ ID 556; ORF 141>:

```
m141.pep
      1  MSFKTDAEIA QSSTMRPIGE IAAKLGLNAD NIEPYGHYKA KINPAEAFKL
     51  PQKQGRLILV TAINPTPAGE GKTTVTIGLA DALRHIGKDA VIALREPSLG
    101  PVFGVKGGAA GGGYAQVLPM EDINLHFTGD FHAIGAANNL LAAMLDNHIY
    151  QGNELNIDPK RVLWRRVVDM NDRQLRNIID GMGKPVDGVM RPDGFDITVA
    201  SEVMAVFCLA KDISDLKERL GNILVAYAKD GSPVYAKDLK ANGAMAALLK
    251  DAIKPNLVQT IEGTPAFVHG GPFANIAHGC NSVTATRLAK HLADYAVTEA
    301  GFGADLGAEK FCDIKCRLAG LKPDAAVVVA TVRALKYNGG VERANLGEEN
    351  LDALEKGLPN LLKHISNLKN VFGLPVVVAL NRFVSDADAE LAMIEKACAE
    401  HGVEVSLTEV WGKGGAGGAD LARKVVNAIE SQTNNFGFAY DVELGIKDKI
    451  RAIAQKVYGA EDVDFSAEAS AEIASLEKLG LDKMPICMAK TQYSLSDNAK
    501  LLGCPEDFRI AVRGITVSAG AGFIVALCGN MMKMPGLPKV PAAEKIDVDA
    551  EGVIHGLF*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 141 shows 97.5% identity over a 558 aa overlap with a predicted ORF (ORF 141.ng) from *N. gonorrhoeae:*

```
m141/g141

                       10        20        30        40        50        60
m141.pep     MSFKTDAEIAQSSTMRPIGEIAAKLGLNADNIEPYGHYKAKINPAEAFKLPQKQGRLILV
             ||||||||| |||||||||||||||||||:||||||||||||||||||||||||||||||
g141         MSFKTDAETAQSSTMRPIGEIAAKLGLNVDNIEPYGHYKAKINPAEAFKLPQKQGRLILV
                       10        20        30        40        50        60
```

```
 201  GGCGGGCGAA GGTAAAACCA CCGTAACCAT CGGTTTGGCG GACGCATTGC
 251  GCCATATCGG CAAAGACTCT GTGATTGCTT TGCGCGAGCC TTCTTTGGGT
 301  CCGGTGTTCG GCGTGAAAGG CGGCGCGGCA GGCGGCGGCT ATGCCCAAGT
 351  TTTGCCGATG GAAGACATCA ACCTGCACTT CACCGGAGAT TTTCACGCCA
 401  TCGGTGCGGC AAATAATCTG CTTGCCGCGA TGCTCGACAA CCATATCTAC
 451  CAAGGCAACG AGTTGAACAT CGACCCCAAA CGCGTGCTGT GGCGGCGCGT
 501  GGTCGATATG AACGACCGCC AGTTGCGCAA CATCATCGAC GGCATGGGCA
 551  AGCCTGTTGA CGGCGTGATG CGTCCTGACG GTTTCGATAT TACCGTTGCT
 601  TCCGAAGTGA TGGCGGTATT CTGTCTTGCC AAAGACATCA GCGATTTGAA
 651  AGAGCGTTTG GGCAACATCC TTGTCGCCTA CGCCAAAGAC GGCAGCCCCG
 701  TTTACGCCAA AGATTTGAAA GCGAATGGCG CGATGGCGGC ATTGCTTAAA
 751  GATGCGATTA AGCCCAACTT GGTGCAAACC ATCGAAGGCA CGCCCGCCTT
 801  CGTACACGGC GGCCCGTTCG CCAACATCGC CCACGGCTGC AACTCCGTAA
 851  CCGCAACCCG TCTGGCGAAA CACCTTGCCG ATTACGCCGT AACCGAAGCA
 901  GGCTTCGGCG CGGACTTGGG CGCGGAAAAA TTCTGCGACA TCAAATGCCG
 951  CCTTGCCGGT TTGAAACCTG ATGCGGCTGT TGTCGTGGCG ACTGTCCGCG
1001  CGTTGAAATA TAACGGCGGC GTGGAACGCG CCAACCTCGG CGAAGAAAAT
1051  TTAGACGCTT TGGAAAAAGG TTTGCCCAAC CTGCTGAAAC ACATTTCCAA
1101  CCTGAAAAAC GTATTCGGAC TGCCCGTCGT CGTTGCGCTC AACCGCTTCG
1151  TGTCCGACTC CGATGCCGAG TTGGCGATGA TTGAAAAAGC CTGTGCCGAA
1201  CACGGCGTTG AAGTTTCCCT GACCGAAGTG TGGGGCAAAG GTGGTGCGGG
1251  CGGCGCGGAT TTGGCGCGCA AAGTCGTCAA [illegible]CCATTGAA AGTCAAACCA
1301  ATAACTTCGG TTTCGCCTAC GATGTCGAGT TGGGCATCAA AGACAAAATC
1351  CGTGCGATTG CCCAAAAAGT GTACGGCGCG GAAGATGTTG ATTTCAGCGC
1401  GGAAGCGTCT GCCGAAATCG CTTCACTGGA AAAACTGGGC TTGGACAAAA
1451  TGCCGATCTG CATGGCGAAA ACCCAATACT CTTTGAGCGA CAACGCCAAA
1501  CTGTTGGGCT GCCCCGAAGA CTTCCGCATC GCCGTGCGCG GCATCACCGT
1551  TTCCGCAGGC GCAGGTTTCA TCGTCGCCCT GTGCGGCAAC ATGATGAAAA
1601  TGCCCGGCCT GCCCAAAGTT CCGGCTGCCG AGAAAATCGA TGTGGACGCA
1651  GAAGGCGTGA TTCACGGCTT GTTCTGA
```

This corresponds to the amino acid sequence <SEQ ID 558; ORF 141.a>:

```
a141.pep
    1  MSFKTDAEIA QSSTMRPIGE IAAKLGLNVD NIEPYGHYKA KINPAEAFKL
   51  PQKQGRLILV TAINPTPAGE GKTTVTIGLA DALRHIGKDS VIALREPSLG
  101  PVFGVKGGAA GGGYAQVLPM EDINLHFTGD FHAIGAANNL LAAMLDNHIY
  151  QGNELNIDPK RVLWRRVVDM NDRQLRNIID GMGKPVDGVM RPDGFDITVA
  201  SEVMAVFCLA KDISDLKERL GNILVAYAKD GSPVYAKDLK ANGAMAALLK
  251  DAIKPNLVQT IEGTPAFVHG GPFANIAHGC NSVTATRLAK HLADYAVTEA
  301  GFGADLGAEK FCDIKCRLAG LKPDAAVVVA TVRALKYNGG VERANLGEEN
  351  LDALEKGLPN LLKHISNLKN VFGLPVVVAL NRFVSDSDAE LAMIEKACAE
  401  HGVEVSLTEV WGKGGAGGAD LARKVVNAIE SQTNNFGFAY DVELGIKDKI
  451  RAIAQKVYGA EDVDFSAEAS AEIASLEKLG LDKMPICMAK TQYSLSDNAK
  501  LLGCPEDFRI AVRGITVSAG AGFIVALCGN MMKMPGLPKV PAAEKIDVDA
  551  EGVIHGLF*
```

m141/a141 99.5% identity in 558 aa overlap

```
                    10        20        30        40        50        60
m141.pep    MSFKTDAEIAQSSTMRPIGEIAAKLGLNADNIEPYGHYKAKINPAEAFKLPQKQGRLILV
            ||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
a141        MSFKTDAEIAQSSTMRPIGEIAAKLGLNVDNIEPYGHYKAKINPAEAFKLPQKQGRLILV
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m141.pep    TAINPTPAGEGKTTVTIGLADALRHIGKDAVIALREPSLGPVFGVKGGAAGGGYAQVLPM
            |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
a141        TAINPTPAGEGKTTVTIGLADALRHIGKDSVIALREPSLGPVFGVKGGAAGGGYAQVLPM
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m141.pep    EDINLHFTGDFHAIGAANNLLAAMLDNHIYQGNELNIDPKRVLWRRVVDMNDRQLRNIID
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a141        EDINLHFTGDFHAIGAANNLLAAMLDNHIYQGNELNIDPKRVLWRRVVDMNDRQLRNIID
                   130       140       150       160       170       180
```

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 561>:

```
m142.seq
      1  ATGCGTGCCG ATTTCATGTT TGCCGACAAT ATGCCCGTGC AGGTGCGCCA
     51  ACGCGCCCTC TATTTCAAGT TGTCCCGTTT TGCCGCGATG CCAGATGTGG
    101  TAGGCAAACC GCTCTTCGGG CGACAGGCCG GTCAGCCCGG CAAAATGTTC
    151  GGCAACATCC TGATGTTCGT CCGCCAGCGT ATTGATGCAG AGGCTGCCGT
    201  TTTCCGACAG GATCGGAATG ATTCGCGCAC TCCGGTTGAT GCACAGCATC
    251  ACGGTCGGCG GCTCGTCGGT AACCGGCGCG ACCGCCGTCA TTGTAATGCC
    301  GTAACGCCCT GCCGCACCGT CTGTCGTGAT GACATGAACG CCTGCCGCGC
    351  AAGATGCCAT CGCATCACGG AACGAAGTTT GAAAATTTTT CTGCAAATCC
    401  GCCATTTTTC CCCTTTAAAC TGTCCCCTAT ATAAGAATGC TGCACACAAG
    451  GCATCCCCcC ATGTGCAGCA GTTTTGA
```

This corresponds to the amino acid sequence <SEQ ID 562; ORF 142>:

```
m142.pep
      1  MRADFMFADN MPVQVRQRAL YFKLSRFAAM PDVVGKPLFG RQAGQPGKMF
     51  GNILMFVRQR IDAEAAVFRQ DRNDSRTPVD AQHHGRRLVG NRRDRRHCNA
    101  VTPCRTVCRD DMNACRARCH RITERSLKIF LQIRHFSPLN CPLYKNAAHK
    151  ASPHVQQF*
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N.gonorrhoeae*</u>

ORF 142 shows 93.7% identity over a 158 aa overlap with a predicted ORF (ORF 142.ng) from *N. gonorrhoeae:*

```
m142/g142

                     10        20        30        40        50        60
m142.pep     MRADFMFADNMPVQVRQRALYFKLSRFAAMPDVVGKPLFGRQAGQPGKMFGNILMFVRQR
             |||||||||||||||||||:||||||||||||::||||||||||||||||||||||||||:
g142         MRADFMFADNMPVQVRQRAFYFKLSRFAAMPNMVGKPLFGRQAGQPGKMFGNILMFVRQH
                     10        20        30        40        50        60

                     70        80        90       100       110       120
m142.pep     IDAEAAVFRQDRNDSRTPVDAQHHGRRLVGNRRDRRHCNAVTPCRTVCRDDMNACRARCH
             ||||||||||||||||||| ||||||||||||:||||||||||||||||||||||||: ||
g142         IDAEAAVFRQDRNDSRTPVYAQHHGRRLVGNRRNRRHCNAVTPCRTVCRDDMNACRTGCH
                     70        80        90       100       110       120

                    130       140       150      159
m142.pep     RITERSLKIFLQIRHFSPLNCPLYKNAAHKASPHVQQFX
             |||||||| |||||||||||| |||||||||||||||||
g142         RITERSLKSFLQIRHFSPLNRPLYKNAAHKASPHVQQFX
                    130       140       150
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 563>:

```
a142.seq
      1  ATGCGTGCCG ATTTCATGTT TGCCGACAAT ATGCCCGTGC AGGTGCGCCA
     51  ACGCGCCCTC TATTTCAAGT TGTCCCGTTT TGCCGCGATG CCAGATGTGG
    101  TAGGCAAACC GCTCTTCGGG CGACAGGCCG GTCAGCCCGG CAAAATGTTC
    151  GGCAACATCC TGATGTTCGT CCGCCAGCGT ATTGATGCAG AGGCTGCCGT
    201  TTTCCGACAG GATCGGAATG ATTCGCGCAC TCCGGTTGAT GCACAGCATC
    251  ACGGTCGGCG GCTCGTCCGT AACCGGCGCA ACCGCCGTCA TTGTAATGCC
    301  GTAACGCCCT GCCGCACCGT CTGTCGTGAT GACATGAACG CCTGCCGCAC
    351  AGGATGCCAT CGCATCACGG AACGAAGTTT GAAAAGTTTT CTGCAAATCC
    401  GCCATTTTTC CCCTTTAAAC TGTCCCCTAT ATAAGAATGC TGCACACAAG
    451  GCACCCCCCA TGTGCAGCAG TTCTGATTCA AAAAGCCGTC GGTCGGACAT
    501  TTCCGCGCGT TACGGCGTAT TACGAGTTCA ACGCATCCTC GATTTTGGCA
    551  AGTTCTGCCA ACAGGTCTTT AAGCAGCAGC ATTTTCTCGC GGCCCAGCAC
    601  TTCCTCGATA GCGTCGTAAC GCTCGTCCAC TTCTTCGCCG ATTTCCTCAT
    651  ACAGCTTCTC GCCCTCGGCA GTCAGCTTCA GAAAAACACG TCGTTGGTCG
    701  TTGGAAGGTT TCAGGCGGAC AACCAAACCC GCTTTTTCAA GGCGGGTCAG
    751  GATACCGGTC AGGCTGGGGC GCAAAATGCA CGCCTGATTC GCCAAATCTT
```

```
1201  CAGGCAACCA TGTTCTTGGT TGCAGGCGCA GTCTTGCTGC TGGGAGCCTT
1251  CTCAGTCTGT CTGATTAAAG AGATCCACGG CGGGGTTTGA
```

This corresponds to the amino acid sequence <SEQ ID 566; ORF 143.ng>:

```
g143.pep
   1  MLSFGYLGVQ TAFTLQSSQM SRIFQTLGAD PHNLGWFFIL PPLAGMLVQP
  51  IVGYYSDRTW KPRLGGRRLP YLLYGTLIAV IVMILMPNSG SFGFGYASLA
 101  ALSFGALMIA LLDVSSNMAM QPFKMMVGDM VNEEQKSYAY GIQSFLANTD
 151  AVVAAILPFV FAYIGLANTA EKGVVPQTVV VAFYVGAALL IITSAFTISK
 201  VKEYDPETYA RYHGIDVAAN QEKANWFELL KTAPKVFWTV TPVQFFCWFA
 251  FRYMWTYSAG AIAENVWHTT DASSVGHQEA GNRYGVLAAV *SVAAVICSF
 301  ILAKVPNKYH KAGYFGCLAL GALGFFSIFF IYNQYALILS YILIGIAWAG
 351  IITYPLTIVA NALSGKHMDT YLGLFNGSVC MPQIVASLLS FVLFPMLGGH
 401  QATMFLVAGA VLLLGAFSVC LIKEIHGGV*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 567>:

```
m143.seq
   1  ATGCTCAGTT TCGGCTTTCT CGGCGTTCAG ACGGCCTTTA CCCTGCAAAG
  51  CTCGCAAATG AGCCGCATTT TTCAAACGCT AGGCGCAGAC CCGCACAATT
 101  TGGGCTGGTT TTTCATCCTG CCGCCGCTGG CGGGGATGCT GGTGCAGCCG
 151  ATTGTCGGCC ATTACTCCGA CCGCACTTGG AAGCCGCGTT TGGGCGGCCG
 201  CCGTCTGCCG TATCTGCTTT ATGGCACGCT GATTGCGGTT ATTGTGATGA
 251  TTTTGATGCC GAACTCGGGC AGCTTCGGTT TCGGCTATGC GTCGCTGGCG
 301  GCTTTGTCGT TCGGCGCGCT GATGATTGCG CTGTTAGACG TGTCGTCAAA
 351  TATGGCGATG CAGCCGTTTA AGATGATGGT CGGCGACATG GTCAACGAGG
 401  AGCAGAAAGG CTACGCCTAC GGGATTCAAA GTTTCTTAGC AAATACGGGC
 451  GCGGTCGTGG CGGCGATTCT GCCGTTTGTG TTTGCGTATA TCGGTTTGGC
 501  GAACACCGCC GAGAAAGGCG TTGTGCCGCA GACCGTGGTC GTGGCGTTTT
 551  ATGTGGGTGC GGCGTTGCTG GTGATTACCA GCGCGTTCAC GATTTTCAAA
 601  GTGAAGGAAT ACGATCCGGA AACCTACGCC CGTTACCACG GCATCGATGT
 651  CGCCCGCGAAT CAGGAAAAAG CCAACTGGAT CGAACTCTTG AAAACCGCGC
 701  CTAAGGCGTT TTGGACGGTT ACTTTGGTGC AATTCTTCTG CTGGTTCGCC
 751  TTCCAATATA TGTGGACTTA CTCGGCAGGC GCGATTGCGG AAAACGTCTG
 801  GCACACCACC GATGCGTCTT CCGTAGGTTA TCAGGAGGCG GGTAACTGGT
 851  ACGGCGTTTT GGCGGCGGTG CAGTCGGTTG CGGCGGTGAT TTGTTCGTTT
 901  GTATTGGCGA AAGTGCCGAA TAAATACCAT AAGGCGGGTT ATTTCGGCTG
 951  TTTGGCTTTG GGCGCGCTCG GCTTTTTCTC CGTTTTCTTC ATCGGCAACC
1001  AATACGCGCT GGTGTTGTCT TATACCTTAA TCGGCATCGC TTGGGCGGGC
1051  ATTATCACTT ATCCGCTGAC GATTGTGACC AACGCCTTGT CGGGCAAGCA
1101  TATGGGCACT TACTTGGGCT TGTTTAACGG CTCTATCTGT ATGCCTCAAA
1151  TCGTCGCTTC GCTGTTGAGT TTCGTGCTTT TCCCTATGCT GGGCGGCTTG
1201  CAGGCCACTA TGTTCTTGGT AGGGGGCGTC GTCCTGCTGC TGGGCGCGTT
1251  TTCCGTGTTC CTGATTAAAG AAACACACGG CGGGGTTTGA
```

This corresponds to the amino acid sequence <SEQ ID 568; ORF 143>:

```
m143.pep
   1  MLSFGFLGVQ TAFTLQSSQM SRIFQTLGAD PHNLGWFFIL PPLAGMLVQP
  51  IVGHYSDRTW KPRLGGRRLP YLLYGTLIAV IVMILMPNSG SFGFGYASLA
 101  ALSFGALMIA LLDVSSNMAM QPFKMMVGDM VNEEQKGYAY GIQSFLANTG
 151  AVVAAILPFV FAYIGLANTA EKGVVPQTVV VAFYVGAALL VITSAFTIFK
 201  VKEYDPETYA RYHGIDVAAN QEKANWIELL KTAPKAFWTV TLVQFFCWFA
 251  FQYMWTYSAG AIAENVWHTT DASSVGYQEA GNWYGVLAAV QSVAAVICSF
 301  VLAKVPNKYH KAGYFGCLAL GALGFFSVFF IGNQYALVLS YTLIGIAWAG
 351  IITYPLTIVT NALSGKHMGT YLGLFNGSIC MPQIVASLLS FVLFPMLGGL
 401  QATMFLVGGV VLLLGAFSVF LIKETHGGV*
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

```
m143 / g143  93.9% identity in 429 aa overlap

                    10        20        30        40        50        60
```

```
 751  TTCCAATATA TGTGGACTTA CTCGGCAGGC GCGATTGCGG AAAACGTCTG
 801  GCACACCACC GATGCGTCTT CCGTAGGTTA TCAGGAGGCG GGTAACTGGT
 851  ACGGCGTTTT GGCGGCGGTG CAGTCGGTTG CGGCGGTGAT TTGTTCGTTT
 901  GTATTGGCGA AAGTGCCGAA TAAATACCAT AAGGCGGGTT ATTTCGGCTG
 951  TTTGGCTTTG GGCGCGCTCG GCTTTTTCTC CGTTTTCTTC ATCGGCAACC
1001  AATACGCGCT GGTGTTGTCT TA-ACCTTAA TCGGCATCGC TTGGGCGGGC
1051  ATTATCACTT ATCCGCTGAC GATTGTGACC AACGCCTTGT CGGGCAAGCA
1101  TATGGGCACT TACTTGGGCC TGTTTAACGG CTCTATCTGT ATGCCGCAAA
1151  TCGTCGCTTC GCTGTTGAGT TTCGTGCTTT TCCCTATGCT GGGCGGCTTG
1201  CAGGCCACTA TGTTCTTGGT AGGGGGCGTC GTCCTGCTGC TGGGCGCGTT
1251  TTCCGTGTTC CTGATTAAAG AAACACACGG CGGGGTTTGA
```

This corresponds to the amino acid sequence <SEQ ID 570; ORF 143.a>:

```
a143.pep
   1  MLSFGFLGVQ TAFTLQSSQM SRIFQTLGAD PHSLGWFFIL PPLAGMLVQP
  51  IVGHYSDRTW KPRLGGRRLP YLLYGTLIAV IVMILMPNSG SFGFGYASLA
 101  ALSFGALMIA LLDVSSNMAM QPFKMMVGDM VNEEQKGYAY GIQSFLANTG
 151  AVVAAILPFV FAYIGLANTA EKGVVPQTVV VAFYVGAALL VITSAFTIFK
 201  VKEYNPETYA RYHGIDVAAN QEKANWIELL KTAPKAFWTV TLVQFFCWFA
 251  FQYMWTYSAG AIAENVWHTT DASSVGYQEA GNWYGVLAAV QSVAAVICSF
 301  VLAKVPNKYH KAGYFGCLAL GALGFFSVFF IGNQYALVLS YTLIGIAWAG
 351  IITYPLTIVT NALSGKHMGT YLGLFNGSIC MPQIVASLLS FVLFPMLGGL
 401  QATMFLVGGV VLLLGAFSVF LIKETHGGV*
```

```
m143/a143    99.5% identity in 429 aa overlap

                    10        20        30        40        50        60
m143.pep     MLSFGFLGVQTAFTLQSSQMSRIFQTLGADPHNLGWFFILPPLAGMLVQPIVGHYSDRTW
             ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
a143         MLSFGFLGVQTAFTLQSSQMSRIFQTLGADPHSLGWFFILPPLAGMLVQPIVGHYSDRTW
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m143.pep     KPRLGGRRLPYLLYGTLIAVIVMILMPNSGSFGFGYASLAALSFGALMIALLDVSSNMAM
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a143         KPRLGGRRLPYLLYGTLIAVIVMILMPNSGSFGFGYASLAALSFGALMIALLDVSSNMAM
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m143.pep     QPFKMMVGDMVNEEQKGYAYGIQSFLANTGAVVAAILPFVFAYIGLANTAEKGVVPQTVV
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a143         QPFKMMVGDMVNEEQKGYAYGIQSFLANTGAVVAAILPFVFAYIGLANTAEKGVVPQTVV
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m143.pep     VAFYVGAALLVITSAFTIFKVKEYDPETYARYHGIDVAANQEKANWIELLKTAPKAFWTV
             ||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||
a143         VAFYVGAALLVITSAFTIFKVKEYNPETYARYHGIDVAANQEKANWIELLKTAPKAFWTV
                   190       200       210       220       230       240

                   250       260       270       280       290       300
m143.pep     TLVQFFCWFAFQYMWTYSAGAIAENVWHTTDASSVGYQEAGNWYGVLAAVQSVAAVICSF
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a143         TLVQFFCWFAFQYMWTYSAGAIAENVWHTTDASSVGYQEAGNWYGVLAAVQSVAAVICSF
                   250       260       270       280       290       300

                   310       320       330       340       350       360
m143.pep     VLAKVPNKYHKAGYFGCLALGALGFFSVFFIGNQYALVLSYTLIGIAWAGIITYPLTIVT
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a143         VLAKVPNKYHKAGYFGCLALGALGFFSVFFIGNQYALVLSYTLIGIAWAGIITYPLTIVT
                   310       320       330       340       350       360

                   370       380       390       400       410       420
m143.pep     NALSGKHMGTYLGLFNGSICMPQIVASLLSFVLFPMLGGLQATMFLVGGVVLLLGAFSVF
```

Homology with a predicted ORF from *N. gonorrhoeae*

```
m144 / g144  91.3% identity in 218 aa overlap

                     10        20        30        40        50        60
m144.pep     MSDTPATRDFGLIDGRAVTGYVLSNRRGTRVCVLDLGGIVQEFSVLADGVRENLVVSFDD
             |||||||||||||||||||||||||||||| | |||||||||||||||||||| ||||||
g144         MSDTPATRDFGLIDGRAVTGYVLSNRRGTCVFVLDLGGIVQEFSVLADGVRENPVVSFDD
                     10        20        30        40        50        60

                     70        80        90       100       110       120
m144.pep     AASYADNPFQINKQIGRVAGRIRGAAFDINGRTYRVEANEGRNALHGGSHGLAVTRFNAV
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g144         AASYADNPFQINKQIGRVAGRIRGAAFDINGRTYRVEANEGRNALHGGSHGLAVTRFNAV
                     70        80        90       100       110       120

                    130       140       150       160       170       180
m144.pep     AADGRSVVLRSRLATVGRRLSQRFGFGYFLPLGRGRPAYRYLSRHRARRHGVRPDAAHLL
             |||             |||||||||  |||||||||||||||||||||||||||||||||
g144         AAD-------------GRRLSQRFG--YFLPLGRGRPAYRYLSRHRARRHGVRPDAAHLL
                                     130       140       150       160

                    190       200       210      219
m144.pep     AAGRGPARCGSAYSAGRTYAGRCRKTARLNGFRRPRSIX
             |||||||||||||||||||:|||||||||||||||||||
g144         AAGRGPARCGSAYSAGRTYSGRCRKTARLNGFRRPRSIX
               170       180       190       200
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 575>:

```
a144.seq
      1  ATGAGCGATA CCCCCGCTAC CCGCGATTTC GGCCTGATCG ACGGGCGTGC
     51  CGTAACCGGC TATGTGCTGT CCAACCGGCG TGGTACGCGT GTCTGCGTGC
    101  TGGACTTGGG CGGGATTGTG CAGGAATTTT CCGTTTTGGC AGACGGCGTG
    151  CGCGAAAACC TCGTGGTGTC GTTCGACGAT GCGGCTTCCT ATGCGGACAA
    201  TCCGTTTCAG ATTAACAAGC AGATAGGGCG CGTGGCCGGA CGCATCCGCG
    251  GTGCGGCGTT CGACATCAAC GGCAGGACTT ACCGCGTGGA GGCCAACGAA
    301  GGCAGGAACG CGCTGCACGG CGGTTCGCAC GGGCTGGCCG TTACCCGTTT
    351  CAACGCGGTG GCGGCAGACG GCCGTTCGGT GGTGCTGCGC AGCCGCCTG.
    401  CAACAGTCGG CCGACGGTTA TCCCAACGAT TTGGATTTGG ATATTTCCTA
    451  CCGCTTGGAC GAGGACGACC GGCTTACCGT TACCTATCGC GCCACCGCGC
    501  TCGGCGACAC GGTGTTCGAC CCGACGCTGC ACATTTACTG GCGGCTGGAC
    551  GCGGGCCTGC ACGATGCGGT TCTGCATATT CCGCAGGGCG GACATATTCC
    601  GGCCGATGCC GAAAAACTGC CCGTCTCAAC GGTTTCAGAC GACCTCGAAG
    651  TATTTGA
```

This corresponds to the amino acid sequence <SEQ ID 576; ORF 144.a>:

```
a144.pep
      1  MSDTPATRDF GLIDGRAVTG YVLSNRRGTR VCVLDLGGIV QEFSVLADGV
     51  RENLVVSFDD AASYADNPFQ INKQIGRVAG RIRGAAFDIN GRTYRVEANE
    101  GRNALHGGSH GLAVTRFNAV AADGRSVVLR SRLXTVGRRL SQRFGFGYFL
    151  PLGRGRPAYR YLSRHRARRH GVRPDAAHLL AAGRGPARCG SAYSAGRTYS
    201  GRCRKTARLN GFRRPRSI*
```

```
m144/a144    99.1% identity in 218 aa overlap

                     10        20        30        40        50        60
m144.pep     MSDTPATRDFGLIDGRAVTGYVLSNRRGTRVCVLDLGGIVQEFSVLADGVRENLVVSFDD
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a144         MSDTPATRDFGLIDGRAVTGYVLSNRRGTRVCVLDLGGIVQEFSVLADGVRENLVVSFDD
                     10        20        30        40        50        60

                     70        80        90       100       110       120
m144.pep     AASYADNPFQINKQIGRVAGRIRGAAFDINGRTYRVEANEGRNALHGGSHGLAVTRFNAV
```

Computer analysis of this amino acid sequence gave the following results:
<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

```
m146 / g146  90.1% identity in 212 aa overlap

                     10        20        30        40        50        60
m146.pep     MAQILLRSRQVVIDHDKVKQYGLLDFMPCLRQPPLDNFPTVRPASVEARGKYVERRRQDK
             | || ||  |||||||||:||||:||||||||||||||||||||  |||||:||||||||
g146         MKQIPLRLLQVVIDHDKVEQYGLFDFMPCLRQPPLDNFPTVRPAPFEARGKHVERRRQDK
                     10        20        30        40        50        60

                     70        80        90       100       110       120
m146.pep     DADGFGQRVANLRRALNVDFQNHVIACRRQRIHTLRACAVIVAKYVGVFQKSFLRDKRLK
             |:|:| |||||||||||||||||||||||||||:|||||||||:|| |||||:|||||:|
g146         DTDSFRQRVANLRRALNVDFQNHVIACRRQRIHALRACAVIVAEYVCVFQKSLLRDKRFK
                     70        80        90       100       110       120

                    130       140       150       160       170       180
m146.pep     LFFGNKVIMYAVCFAFTRRARRVRHGNAQTVMVCQQPRHQRGFARAGSGRNDKDVAFSIS
             ||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
g146         LFFGNKVIMYAVCFAFTRRARRMRHGNAQTVMVCQQPRHQRGFARAGSGRNDKDVAFSIS
                    130       140       150       160       170       180

                    190       200       210
m146.pep     GHIFYLYIFQPIVSQWTPSFLFADAHILPLLFX
             ||||||||||||||| || |:||||||||||||
g146         GHIFYLYIFQPIVSQRTPYFIFADAHILPLLFX
                    190       200       210
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 581>:

```
a146.seq
      1  ATGGCGCAAA TCCTCCTCCG CCCGCGCCAA GTCATCATTG ACCACGACAA
     51  AATCGAACAA TACGGACTGT TCGATTTCAT GCCTTGCCTT CGACAGCCTC
    101  CTTTGGATAA CTTCCCGACT GTCCGTCCCG CGTCCGTTGA GACGCGCAGC
    151  AAGCACATCG AAAGACGGCG GCAGGATAAA GATGCCGACG GCTTCGGGCA
    201  GCGCATCTCG AACCTGAGCC GCGCCCTGAA CGTCGATTTC CAAAATCACG
    251  TCATAACCTG CCGCCGCCAA CGCATTCACA CCCTCCGCGC TTGTGCCGTA
    301  ATAGTTGCCG AACACGTCCG CGTATTCCAA AAAAGCCTCC TGCGCGATAA
    351  GCGACTCAAA CTCTTCTTTG GAAACAAAGT GATAATGTAC GCCGTTTGCT
    401  TCGCCTTCAC GCGGCGGACG CGTCGTGTGC GACACGGAAA CGCGCAAACC
    451  GTTATGGTTT GCCAACAGCC GCGACACCAG CGTGGTTTTG CCCGTGCCGG
    501  AAGCGGCCGA AATGATAAAG ATGTTGCCTT TTCGATAAGC GGACATATTT
    551  TTTACCTGTA TATTTTCCAG CCGATTGTAT CACAACGGAC ACCCGGTTTC
    601  CTATTTGCCG ATGCCCATAT TTTGCCGCTA TTGTTTTGA
```

This corresponds to the amino acid sequence <SEQ ID 582; ORF 146.a>:

```
a146.pep
      1  MAQILLRPRQ VIIDHDKIEQ YGLFDFMPCL RQPPLDNFPT VRPASVETRS
     51  KHIERRRQDK DADGFGQRIS NLSRALNVDF QNHVITCRRQ RIHTLRACAV
    101  IVAEHVRVFQ KSLLRDKRLK LFFGNKVIMY AVCFAFTRRT RRVRHGNAQT
    151  VMVCQQPRHQ RGFARAGSGR NDKDVAFSIS GHIFYLYIFQ PIVSQRTPGF
    201  LFADAHILPL LF*
```

```
m146/a146     90.6% identity in 212 aa overlap

                     10        20        30        40        50        60
m146.pep     MAQILLRSRQVVIDHDKVKQYGLLDFMPCLRQPPLDNFPTVRPASVEARGKYVERRRQDK
             ||||||| |||:|||||::||||:|||||||||||||||||||||||:|:|::|||||||
a146         MAQILLRPRQVIIDHDKIEQYGLFDFMPCLRQPPLDNFPTVRPASVETRSKHIERRRQDK
                     10        20        30        40        50        60

                     70        80        90       100       110       120
```

```
1101   AGGTCGTCTG AAAGGCAGCT GGGGCGTGCA ATATTTACAA CAAAAATCCA
1151   GTGCTTTATC TGCCATATCC GAAGCGGTTA AACAACCGAT GCTGCTTGAC
1201   AACAAAGTGC AACATTACAG CTTTTTCGGT GTAGAACAGG CAAACTGGGA
1251   CAACTTCACG CTTGAAGGAG GCGTACGCGT GGAAAAACAA AAAGCCTCCA
1301   TTCAGTACGA CAAAGCATTG ATTGATCGGG AAAACTACTA CAACCACCCC
1351   CTGCCCGACC TCGGCGCGCA CCGCCAAACC GCCCGCTCAT TCGCACTTTC
1401   GGGCAACTGG TATTTCACGC CACAACACAA ACTCAGCCTG ACCGCCTCCC
1451   ATCAGGAACG CCTGCCGTCA ACGCAAGAGC TGTACGCACA CGGCAAACAC
1501   GTCGCCACCA ACACCTTTGA AGTCGGCAAC AAACACCTCA ACAAAGAGCG
1551   TTCCAACAAT ATCGAACTCG CGCTGGGCTA CGAAGGCGAC CGCTGGCAAT
1601   ACAATCTGGC ACTCTACCGC AACCGCTTCG GTAACTACAT TTACGCCCAA
1651   ACCTTAAACG ACGGACGCGG CCCCAAATCC ATCGAAGACG ACAGCGAAAT
1701   GAAGCTCGTG CGCTACAACC AATCCGGCGC CGACTTCTAC GGCGCGGAAG
1751   GCGAAATCTA CTTCAAACCG ACACCGCGCT ACCGCATCGG CGTTTCCGGC
1801   GACTATGTAC GAGGCCGTCT GAAAAACCTG CCTTCCCTAC CCGGCAGAGA
1851   AGATGCCTAC GGCAACCGTC CTTTCATCGC ACAGGACGAC CAAAATGCCC
1901   CCCGTGTTCC GGCTGCGCGC CTCGGCTTCC ACCTGAAAGC CTCGCTGACC
1951   GACCGTATCG ATGCCAATTT GGACTACTAC CGCGTGTTCG CCCAAAACAA
2001   ACTCGCCCGC TACGAAACGC GCACGCCCGG ACACCATATG CTCAACCTCG
2051   GCGCAAACTA CCGCCGCAAT ACGCGCTATG GCGAGTGGAA TTGGTACGTC
2101   AAAGCCGACA ACCTGCTCAA CCAATCCGTT TACGCCCACA GCAGCTTTCT
2151   CTCTGATACG CCGCAAATGG GCCGCAGCTT TACCGGCGGC GTGAACGTGA
2201   AGTTTTAA
```

This corresponds to the amino acid sequence <SEQ ID 586; ORF 147>:

```
m147.pep  (partial)
  1   ..PHKTEQSVDL ETVSVVGKSR PRATSGLLHT STASDKIISG DTLRQKAVNL
 51     GDALDGVPGI HASQYGGGAS APVIRGQTGR RIKVLNHHGE TGDMADFSPD
101     HAIMVDTALS QQVEILRGPV TLLYSSGNVA GLVDVADGKI PEKMPENGVS
151     GELGLRLSSG NLEKLTSGGI NIGLGKNFVL HTEGLYRKSG DYAVPRYRNL
201     KRLPDSHADS QTGSIGLSWV GEKGFIGVAY SDRRDQYGLP AHSHEYDDCH
251     ADIIWQKSLI NKRYLQLYPH LLTEEDIDYD NPGLSCGFHD DDNAHAHTHS
301     GRPWIDLRNK RYELRAEWKQ PFPGFEALRV HLNRNDYRHD EKAGDAVENF
351     FNNQTQNARI ELRHQPIGRL KGSWGVQYLQ QKSSALSAIS EAVKQPMLLD
401     NKVQHYSFFG VEQANWDNFT LEGGVRVEKQ KASIQYDKAL IDRENYYNHP
451     LPDLGAHRQT ARSFALSGNW YFTPQHKLSL TASHQERLPS TQELYAHGKH
501     VATNTFEVGN KHLNKERSNN IELALGYEGD RWQYNLALYR NRFGNYIYAQ
551     TLNDGRGPKS IEDDSEMKLV RYNQSGADFY GAEGEIYFKP TPRYRIGVSG
601     DYVRGRLKNL PSLPGREDAY GNRPFIAQDD QNAPRVPAAR LGFHLKASLT
651     DRIDANLDYY RVFAQNKLAR YETRTPGHHM LNLGANYRRN TRYGEWNWYV
701     KADNLLNQSV YAHSSFLSDT PQMGRSFTGG VNVKF*
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

```
m147 / g147  92.3% identity in 142 aa overlap

                                            10        20        30
m147.pep                             PHKTEQSVDLETVSVVGKSRPRATSGLLHTS
                                      |:||||| ||||||||||||||||||||||
g147        MRREAKMAQITLKPIVLSILLINTPLLAQAHETEQSVGLETVSVVGKSRPRATSGLLHTS
                    10        20        30        40        50        60

                    40        50        60        70        80        90
m147.pep    TASDKIISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGET
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g147        TASDKIISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGET
                    70        80        90       100       110       120

                   100       110       120       130       140       150
m147.pep    GDMADFSPDHAIMVDTALSQQVEILRGPVTLLYSSGNVAGLVDVADGKIPEKMPENGVSG
            ||||||||||||||||||||||||||||||||||||||||  :    | | |
g147        GDMADFSPDHAIMVDTALSQQVEILRGPVTLLYSSGNVAGAGQCCRWKNPPKNA
```

```
601  YNQSGADFYG AEGEIYFKPT PRYRIGVSGD YVRGRLKNLP SLPGREDAYG
651  NRPLIAQADQ NAPRVPAARL GVHLKASLTD RIDANLDYYR VFAQNKLARY
701  ETRTPGHHML NLGANYRRNT RYGEWNWYVK ADNLLNQSVY AHSSFLSDTP
751  QMGRSFTGGV NVKF*
```

```
m147/a147     98.1% identity in 734 aa overlap

                                                10        20        30
m147.pep                                PHKTEQSVDLETVSVVGKSRPRATSGLLHTS
                                         |  |||| |||||||||||||||||||||
a147          MRREAKMAQTTLKPIVLSILLINTPLLSQAHGTEQSVGLETVSVVGKSRPRATSGLLHTS
                      10        20        30        40        50        60

                      40        50        60        70        80        90
m147.pep      TASDKIISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGET
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a147          TASDKIISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGET
                      70        80        90       100       110       120

                     100       110       120       130       140       150
m147.pep      GDMADFSPDHAIMVDTALSQQVEILRGPVTLLYSSGNVAGLVDVADGKIPEKMPENGVSG
              |||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
a147          GDMADFSPDHAIMVDSALSQQVEILRGPVTLLYSSGNVAGLVDVADGKIPEKMPENGVSG
                     130       140       150       160       170       180

                     160       170       180       190       200       210
m147.pep      ELGLRLSSGNLEKLTSGGINIGLGKNFVLHTEGLYRKSGDYAVPRYRNLKRLPDSHADSQ
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a147          ELGLRLSSGNLEKLTSGGINIGLGKNFVLHTEGLYRKSGDYAVPRYRNLKRLPDSHADSQ
                     190       200       210       220       230       240

                     220       230       240       250       260       270
m147.pep      TGSIGLSWVGEKGFIGVAYSDRRDQYGLPAHSHEYDDCHADIIWQKSLINKRYLQLYPHL
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
a147          TGSIGLSWVGEKGFIGAAYSDRRDQYGLPAHSHEYDDCHADIIWQKSLINKRYLQLYPHL
                     250       260       270       280       290       300

                     280       290       300       310       320       330
m147.pep      LTEEDIDYDNPGLSCGFHDDDNAHAHTHSGRPWIDLRNKRYELRAEWKQPFPGFEALRVH
              |||||||||||||||||||||:||||:|:|:|||||||||||||||||||||||||||||
a147          LTEEDIDYDNPGLSCGFHDDDAHAHAHNGKPWIDLRNKRYELRAEWKQPFPGFEALRVH
                     310       320       330       340       350       360

                     340       350       360       370       380       390
m147.pep      LNRNDYRHDEKAGDAVENFFNNQTQNARIELRHQPIGRLKGSWGVQYLQQKSSALSAISE
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||| ||
a147          LNRNDYRHDEKAGDAVENFFNNQTQNARIELRHQPIGRLKGSWGVQYLGQKSSALSATSE
                     370       380       390       400       410       420

                     400       410       420       430       440       450
m147.pep      AVKQPMLLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASIQYDKALIDRENYYNHPL
              |||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
a147          AVKQPMLLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASIRYDKALIDRENYYNHPL
                     430       440       450       460       470       480

                     460       470       480       490       500       510
m147.pep      PDLGAHRQTARSFALSGNWYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNK
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a147          PDLGAHRQTARSFALSGNWYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNK
                     490       500       510       520       530       540

                     520       530       540       550       560       570
m147.pep      HLNKERSNNIELALGYEGDRWQYNLALYRNRFGNYIYAQTLNDGRGPKSIEDDSEMKLVR
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a147          HLNKERSNNIELALGYEGDRWQYNLALYRNRFGNYIYAQTLNDGRGPKSIEDDSEMKLVR
```

```
101  KGKLPFETVS QSYALEYGEA AVEIHTDAVK LGSRVLLVDD LIATGGTMLA
151  GLELIRKLGG EIVEAAAILE FTDLQGGKNI RASGAPLFTL LQNEGCMKG*
```

Computer analysis of this amino acid sequence gave the following results:
<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

```
m148 / g148   99.0% identity in 199 aa overlap

                       10        20        30        40        50        60
m148.pep      MALKTSNLEHAMLVHPEAMSVGALADKIRKIENWPQKGILFHDITPVLQSAEYFRLLVDL
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g148          MALKTSNLEHAMLVHPEAMSVGALADKIRKIENWPQKGILFHDITPVLQSAEYFRLLVDL
                       10        20        30        40        50        60

                       70        80        90       100       110       120
m148.pep      LVYRYMDQKIDIVAGLDARGFIIGAALAYQLNVGFVPIRKKGKLPFETVSQSYALEYGEA
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g148          LVYRYMDQKIDIVAGLDARGFIIGAALAYQLNVGFVPIRKKGKLPFETVSQSYALEYGEA
                       70        80        90       100       110       120

                      130       140       150       160       170       180
m148.pep      AVEIHTDAVKLGSRVLLVDDLIATGGTMLAGLELIRKLGGEIVEAAAILEFTDLQGGKNI
              |||||||||| |||||||||:||||||||||||||||||||||||||||||||||||||
g148          AVEIHTDAVKPGSRVLLVDDLVATGGTMLAGLELIRKLGGEIVEAAAILEFTDLQGGKNI
                      130       140       150       160       170       180

                      190       200
m148.pep      RASGAPLFTLLQNEGCMKGX
              |||||||||||||||||||
g148          RASGAPLFTLLQNEGCMKGX
                      190       200
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 593>:

```
a148.seq
      1  ATGGCGTTAA AAACATCAAA CTTGGAACAC GCAATGCTGG TTCATCCCGA
     51  AGCTATGAGT GTCGGTGCGC TTGCCGACAA AATCCGCAAA ATCGAAAACT
    101  GGCCGCAAAA AGGCATCTTA TTCCACGACA TCACGCCCGT CCTGCAAAGC
    151  GCGGAATACT TCCGACTTTT GGTTGATTTA TTGGTTTACC GCTATATGGA
    201  TCAGAAAATC GACATCGTTG CCGGTTTGGA CGCGCGCGGC TTCATTATCG
    251  GCGCGGCACT CGCCTACCAG CTCAACGTCG GTTTCGTCCC CATCCGCAAA
    301  AAAGGCAAGC TGCCTTTTGA AACCGTATCG CAAAGCTACG CGCTCGAATA
    351  CGGGGAAGCT GCGGTGGAAA TCCACACCGA TGCCGTCAAA CTCGGTTCGC
    401  GCGTGCTGCT GGTCGATGAT TTGGTTGCCA CGGGCGGCAC GATGCTTGCC
    451  GGACTGGAGC TGATCCGCAA ACTCGGCGGG GAAATTGTCG AAGCCGCCGC
    501  CATTTTGGAA TTTACCGACC TTCAAGGCGG CAAGAATATC CGTGCAAGCG
    551  GCGCGCCCTT ATTTACCCTG CTTCAAAACG AAGGCTGTAT GAAGGGCTGA
```

This corresponds to the amino acid sequence <SEQ ID 594; ORF 148.a>:

```
a148.pep
      1  MALKTSNLEH AMLVHPEAMS VGALADKIRK IENWPQKGIL FHDITPVLQS
     51  AEYFRLLVDL LVYRYMDQKI DIVAGLDARG FIIGAALAYQ LNVGFVPIRK
    101  KGKLPFETVS QSYALEYGEA AVEIHTDAVK LGSRVLLVDD LVATGGTMLA
    151  GLELIRKLGG EIVEAAAILE FTDLQGGKNI RASGAPLFTL LQNEGCMKG*
```

```
m148/a148     99.5% identity in 199 aa overlap

                       10        20        30        40        50        60
m148.pep      MALKTSNLEHAMLVHPEAMSVGALADKIRKIENWPQKGILFHDITPVLQSAEYFRLLVDL
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a148          MALKTSNLEHAMLVHPEAMSVGALADKIRKIENWPQKGILFHDITPVLQSAEYFRLLVDL
                       10        20        30        40        50        60

                       70        80        90       100       110       120
```

```
 451  ATTTACGCCC AAACCTTAAA CGACGGACGC GGCCCCAAAT CCATCGAAGA
 501  CGACAGCGAA ATGAAGCTCG TGCGCTACAA CCAATCCGGC GCCGACTTCT
 551  ACGGCGCGGA AGGCGAAATC TACTTCAAAC CGACACCGCG CTACCGCATC
 601  GGCGTTTCCG GCGACTATGT ACGAGGCCGT CTGAAAAACC TGCCTTCCCT
 651  ACCCGGCAGA GAAGATGCCT ACGGCAACCG TCCTTTCATC GCACAGGACG
 701  ACCAAAATGC CCCCCGTGTT CCGGCTGCGC GCCTCGGCTT CCACCTGAAA
 751  GCCTCGCTGA CCGACCGTAT CGATGCCAAT TTGGACTACT ACCGCGTGTT
 801  CGCCCAAAAC AAACTCGCCC GCTACGAAAC GCGCACGCCC GGACACCATA
 851  TGCTCAACCT CGGCGCAAAC TACCGCCGCA ATACGCGCTA TGGCGAGTGG
 901  AATTGGTACG TCAAAGCCGA CAACCTGCTC AACCAATCCG TTTACGCCCA
 951  CAGCAGCTTT CTCTCTGATA CGCCGCAAAT GGGCCGCAGC TTTACCGGCG
1001  GCGTGAACGT GAAGTTTTAA
```

This corresponds to the amino acid sequence <SEQ ID 598; ORF 149>:

```
m149.pep
  1  MLLDNKVQHY SFFGVEQANW DNFTLEGGVR VEKQKASIQY DKALIDRENY
 51  YNHPLPDLGA HRQTARSFAL SGNWYFTPQH KLSLTASHQE RLPSTQELYA
101  HGKHVATNTF EVGNKHLNKE RSNNIELALG YEGDRWQYNL ALYRNRFGNY
151  IYAQTLNDGR GPKSIEDDSE MKLVRYNQSG ADFYGAEGEI YFKPTPRYRI
201  GVSGDYVRGR LKNLPSLPGR EDAYGNRPFI AQDDQNAPRV PAARLGFHLK
251  ASLTDRIDAN LDYYRVFAQN KLARYETRTP GHHMLNLGAN YRRNTRYGEW
301  NWYVKADNLL NQSVYAHSSF LSDTPQMGRS FTGGVNVKF*
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N.gonorrhoeae*</u>

ORF 149 shows 95.9% identity over a 339 aa overlap with a predicted ORF (ORF 149.ng) from *N. gonorrhoeae:*

```
m149/g149

                   10        20        30        40        50        60
m149.pep   MLLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASIQYDKALIDRENYYNHPLPDLGA
           ||:||:|:||||||||||||||||||||||||||||||:||||||||||||:||||||
g149       MLIDNNVRHYSFFGVEQANWDNFTLEGGVRVEKQKASIRYDKALIDRENYYNQPLPDLGA
                   10        20        30        40        50        60

                   70        80        90       100       110       120
m149.pep   HRQTARSFALSGNWYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKE
           |||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
g149       HRQTARSFALSGNWYFTPHHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKE
                   70        80        90       100       110       120

                  130       140       150       160       170       180
m149.pep   RSNNIELALGYEGDRWQYNLALYRNRFGNYIYAQTLNDGRGPKSIEDDSEMKLVRYNQSG
           |||||||||||:||||||||| ||||||||||||||||||||||||||||||||||||||
g149       RSNNIELALGYKGDRWQYNLAAYRNRFGNYIYAQTLNDGRGPKSIEDDSEMKLVRYNQSG
                  130       140       150       160       170       180

                  190       200       210       220       230       240
m149.pep   ADFYGAEGEIYFKPTPRYRIGVSGDYVRGRLKNLPSLPGREDAYGNRPFIAQDDQNAPRV
           |||||||||||||||||||||||||||||||||||||||||| ||:|||||| |||||:
g149       ADFYGAEGEIYFKPTPRYRIGVSGDYVRGRLKNLPSLPGREDPYGKRPFIAQADQNAPRI
                  190       200       210       220       230       240

                  250       260       270       280       290       300
m149.pep   PAARLGFHLKASLTDRIDANLDYYRVFAQNKLARYETRTPGHHMLNLGANYRRNTRYGEW
           ||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
g149       PAARLGFHLKTSLTDRIDANLDYYRVFAQNKLARYETRTPGHHMLNLGANYRRNTRYGEW
                  250       260       270       280       290       300

                  310       320       330       340
m149.pep   NWYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKFX
           ||||||||||||||||||||||||||||||:||||||||
```

```
                 |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a149             PAARLGVHLKASLTDRIDANLDYYRVFAQNKLARYETRTPGHHMLNLGANYRRNTRYGEW
                        250       260       270       280       290       300

                        310       320       330       340
m149.pep         NWYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKFX
                 ||||||||||||||||||||||||||||||||||||||||
a149             NWYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKFX
                        310       320       330       340
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 601>:

g149-1.seq

```
   1  ATGGCACAAA TCACACTCAA ACCCATTGTT TTATCAATTC TTTTAATCAA
  51  CACACCCCTC CTCGCCCAAG CGCATGAAAC TGAGCAATCG GTGGGCTTGG
 101  AAACGGTCAG CGTCGTCGGC AAAAGCCGTC CGCGCGCGAC TTCGGGGCTG
 151  CTGCACACTT CGACCGCCTC CGACAAAATC ATCTCCGGCG ATACTTTGCG
 201  CCAAAAAGCC GTCAACTTGG GCGACGCTTT GGACGGCGTA CCGGGCATCC
 251  ACGCTTCGCA ATACGGCGGC GGCGCATCCG CTCCCGTTAT TCGCGGTCAA
 301  ACGGGCAGAC GGATTAAAGT ATTGAACCAT CACGGCGAAA CGGGCGATAT
 351  GGCGGACTTT TCTCCCGAT[illegible] ACGCCATTAT GGTAGATACC GCCTTGTCGC
 401  AACAGGTTGA AATCCTGCGC GGGCCGGTTA CGCTCTTGTA CAGCTCGGGC
 451  AATGTGGCGG GGCTGGTCGA TGTTGCCGAT GGAAAAATCC CCGAAAAAAT
 501  GCCTGAAAAC GGCGTATCGG GCGaagccgG ATTGCGTTTG AGCAGCGGCA
 551  ATTTAGAAAA ACTGACATCC GCAGGCATCA ATATCGGACT GGGCAAAAAC
 601  TTCGTGCTGC ATACCGAAGG CTTGTACCGC AAATCGGGCG ATTACGCCGT
 651  ACCGCGTTAC CGCAATCTGA AACGCCTGCC CGACAGCCAT GCCGATTCGC
 701  AAACGGGCAG CATCGGGCTG TCTTGGGTGG GCGAAAAAGG CTTTATCGGC
 751  GCAGCATACA GCGACCGTCG CGACCGCTAC GGCCTGCCTG CCCACAGCCA
 801  CGAATACGAT GATTGCCACG CCGACATCAT CTGGCAAAAG AGTTTGATCA
 851  ACAAACGCTA TTTGCAGCTT TATCCGCACT TGTTGACCGA AGAAGACATC
 901  GATTACGACA ATCCGGGCTT GAGCTGCGGC TTCCACGACG GCGACGGTGC
 951  ACACGCACAC ACCCACAACG GCAAACCGTG GATAGACCTG CGCAACAAAC
1001  GCTACGAACT CCGCGCCGAA TGGAAGCAGC CATTCCCCGG TTTTGAAGCC
1051  CTGCGCGTAC ATCTGAACCG CAATGACTAC CACCACGACG AAAAAGCAGG
1101  CGATGCAGTA GAAAACTTCT TCAACAACAA AACACACAAC GCCCGTATCG
1151  AGTTGCGCCA CCAACCCATA GGCCGTCTGA AAGGCAGCTG GGGCGTGCAA
1201  TATTTGGGAC AAAAATCCAG CGCGCTTTCC GCCATTCCCG AAACCGTCCA
1251  ACAACCGATG TTGATTGACA ACAATGTCCG CCATTACAGC TTTTTCGGTG
1301  TAGAACAGGC AAATTGGGAC AACTTCACGC TTGAAGGCGG CGTACGCGTG
1351  GAAAAACAAA AAGCCTCCAT CCGGTACGAC AAAGCATTGA TTGATCGAGA
1401  AAACTACTAC AACCAGCCCC TGCCCGACCT CGGCGCGCAC CGCCAAACCG
1451  CCCGCTCGTT CGCACTTTCG GGCAACTGGT ATTTCACGCC ACACCACAAA
1501  CTCAGCCTGA CCGCCTCCCA TCAGGAACGC CTGCCGTCAA CGCAAGAACT
1551  GTACGCACAC GGCAAGCACG TCGCCACCAA CACCTTTGAA GTCGGCAACA
1601  AACACCTCAA CAAAGAGCGT TCCAACAATA TCGAACTCGC GCTGGGCTAC
1651  GAAGGCGACC GCTGGCAATA CAATCTGGCA GCCTACCGCA ACCGATTCGG
1701  CAACTACATT TACGCCCAAA CCTTAAACGA CGGACGCGGC CCCAAATCCA
1751  TCGAAGACGA CAGCGAAATG AAGCTCGTGC GCTACAACCA ATCCGGTGCC
1801  GACTTCTACG GCGCGGAAGG CGAAATCTAC TTCAAACCGA CACCGCGCTA
1851  CCGCATCGGT GTTTCCGGCG ACTATGTACG AGGCCGTCTG AAAAACCTGC
1901  CGTCCCTACC CGGCAGGGAA GATCCCTACG GCAAACGTCC CTTCATCGCA
1951  CAAGCCGACC AAAACGCCCC CCGCATTCCG GCTGCGCGCC TCGGCTTCCA
2001  CCTGAAAACC TCGCTAACCG ACCGTATCGA TGCCAATTTG GACTACTACC
2051  GCGTGTTCGC CCAAAACAAA CTCGCCCGCT ACGAAACGCG TACGCCCGGA
2101  CACCATATGC TCAACCTCGG TGCAAACTAC CGCCGCAATA CGCGCTATGG
2151  CGAGTGGAAT TGGTACGTCA AAGCCGACAA CCTGCTCAAC CAATCCGTTT
2201  ACGCCCACAG CAGCTTCCTC TCTGATACGC CGCAAATGGG CCGCAGCTTT
2251  ACCGGCGGCG TAAACGTGAA GTTTTAA
```

This corresponds to the amino acid sequence <SEQ ID 602; ORF 149-1.ng>:

g149-1.pep

```
   1  MAQITLKPIV LSILLINTPL LAQAHETEQS VGLETVSVVG KSRPRATSGL
  51  LHTSTASDKI ISGDTLRQKA VNLGDALDGV PGIHASQYGG GASAPVIRGQ
 101  TGRRIKVLNH HGETGDMADF SPDHAIMVDT ALSQQVEILR GPVTLLYSSG
 151  NVAGLVDVAD GKIPEKMPEN GVSGEAGLRL SSGNLEKLTS AGINIGLGKN
 201  FVLHTEGLYR KSGDYAVPRY RNLKRLPDSH ADSQTGSIGL SWVGEKGFIG
 251  AAYSDRRDRY GLPAHSHEYD DCHADIIWQK SLINKRYLQL YPHLLTEEDI
 301  DYDNPGLSCG FHDGDGAHAH THNGKPWIDL RNKRYELRAE WKQPFPGFEA
```

```
551  EGDRWQYNLA LYRNRFGNYI YAQTLNDGRG PKSIEDDSEM KLVRYNQSGA
601  DFYGAEGEIY FKPTPRYRIG VSGDYVRGRL KNLPSLPGRE DAYGNRPFIA
651  QDDQNAPRVP AARLGFHLKA SLTDRIDANL DYYRVFAQNK LARYETRTPG
701  HHMLNLGANY RRNTRYGEWN WYVKADNLLN QSVYAHSSFL SDTPQMGRSF
751  TGGVNVKF*
```

m149-1/g149-1   96.2% identity in 758 aa overlap

```
                       10        20        30        40        50        60
m149-1.pep     MAQTTLKPIVLSILLINTPLLAQAHETEQSVDLETVSVVGKSRPRATSGLLHTSTASDKI
               ||| ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
g149-1         MAQITLKPIVLSILLINTPLLAQAHETEQSVGLETVSVVGKSRPRATSGLLHTSTASDKI
                       10        20        30        40        50        60

                       70        80        90       100       110       120
m149-1.pep     ISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGETGDMADF
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g149-1         ISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGETGDMADF
                       70        80        90       100       110       120

                      130       140       150       160       170       180
m149-1.pep     SPDHAIMVDTALSQQVEILRGPVTLLYSSGNVAGLVDVADGKIPEKMPENGVSGELGLRL
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
g149-1         SPDHAIMVDTALSQQVEILRGPVTLLYSSGNVAGLVDVADGKIPEKMPENGVSGEAGLRL
                      130       140       150       160       170       180

                      190       200       210       220       230       240
m149-1.pep     SSGNLEKLTSGGINIGLGKNFVLHTEGLYRKSGDYAVPRYRNLKRLPDSHADSQTGSIGL
               ||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
g149-1         SSGNLEKLTSAGINIGLGKNFVLHTEGLYRKSGDYAVPRYRNLKRLPDSHADSQTGSIGL
                      190       200       210       220       230       240

                      250       260       270       280       290       300
m149-1.pep     SWVGEKGFIGVAYSDRRDQYGLPAHSHEYDDCHADIIWQKSLINKRYLQLYPHLLTEEDI
               ||||||||||:|||||||:|||||||||||||||||||||||||||||||||||||||||
g149-1         SWVGEKGFIGAAYSDRRDRYGLPAHSHEYDDCHADIIWQKSLINKRYLQLYPHLLTEEDI
                      250       260       270       280       290       300

                      310       320       330       340       350       360
m149-1.pep     DYDNPGLSCGFHDDDNAHAHTHSGRPWIDLRNKRYELRAEWKQPFPGFEALRVHLNRNDY
               ||||||||||||| |:||||||:|:|||||||||||||||||||||||||||||||||||
g149-1         DYDNPGLSCGFHDGDGAHAHTHNGKPWIDLRNKRYELRAEWKQPFPGFEALRVHLNRNDY
                      310       320       330       340       350       360

                      370       380       390       400       410       420
m149-1.pep     RHDEKAGDAVENFFNNQTQNARIELRHQPIGRLKGSWGVQYLQQKSSALSAISEAVKQPM
               :|||||||||||||||:|:||||||||||||||||||||||| ||||||||| |:|:|||
g149-1         HHDEKAGDAVENFFNNKTHNARIELRHQPIGRLKGSWGVQYLGQKSSALSAIPETVQQPM
                      370       380       390       400       410       420

                      430       440       450       460       470       480
m149-1.pep     LLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASIQYDKALIDRENYYNHPLPDLGAH
               |:||:|:||||||||||||||||||||||||||||||:|||||||||||||:||||||||
g149-1         LIDNNVRHYSFFGVEQANWDNFTLEGGVRVEKQKASIRYDKALIDRENYYNQPLPDLGAH
                      430       440       450       460       470       480

                      490       500       510       520       530       540
m149-1.pep     RQTARSFALSGNWYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKER
               |||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
g149-1         RQTARSFALSGNWYFTPHHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKER
                      490       500       510       520       530       540

                      550       560       570       580       590       600
m149-1.pep     SNNIELALGYEGDRWQYNLALYRNRFGNYIYAQTLNDGRGPKSIEDDSEMKLVRYNQSGA
               |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
g149-1         SNNIELALGYEGDRWQYNLAAYRNRFGNYIYAQTLNDGRGPKSIEDDSEMKLVRYNQSGA
                      550       560       570       580       590       600

                      610       620       630       640       650       660
m149-1.pep     DFYGAEGEIYFKPTPRYRIGVSGDYVRGRLKNLPSLPGREDAYGNRPFIAQDDQNAPRVP
               ||||||||||||||||||||||||||||||||||||||||| ||:|||||| ||||||:|
g149-1         DFYGAEGEIYFKPTPRYRIGVSGDYVRGRLKNLPSLPGREDPYGKRPFIAQADQNAPRIP
                      610       620       630       640       650       660
```

```
351  LRVHLNRNDY RHDEKAGDAV ENFFNNQTQN ARIELRHQPI GRLKGSWGVQ
401  YLGQKSSALS ATSEAVKQPM LLDNKVQHYS FFGVEQANWD NFTLEGGVRV
451  EKQKASIRYD KALIDRENYY NHPLPDLGAH RQTARSFALS GNWYFTPQHK
501  LSLTASHQER LPSTQELYAH GKHVATNTFE VGNKHLNKER SNNIELALGY
551  EGDRWQYNLA LYRNRFGNYI YAQTLNDGRG PKSIEDDSEM KLVRYNQSGA
601  DFYGAEGEIY FKPTPRYRIG VSGDYVRGRL KNLPSLPGRE DAYGNRPLIA
651  QADQNAPRVP AARLGVHLKA SLTDRIDANL DYYRVFAQNK LARYETRTPG
701  HHMLNLGANY RRNTRYGEWN WYVKADNLLN QSVYAHSSFL SDTPQMGRSF
751  TGGVNVKF*
```

a149-1/m149-1   98.0% identity in 758 aa overlap

```
                          10        20        30        40        50        60
a149-1.pep        MAQTTLKPIVLSILLINTPLLSQAHGTEQSVGLETVSVVGKSRPRATSGLLHTSTASDKI
                  ||||||||||||||||||||:|||| ||||| ||||||||||||||||||||||||||||
m149-1            MAQTTLKPIVLSILLINTPLLAQAHETEQSVDLETVSVVGKSRPRATSGLLHTSTASDKI
                          10        20        30        40        50        60

                          70        80        90       100       110       120
a149-1.pep        ISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGETGDMADF
                  |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
m149-1            ISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGETGDMADF
                          70        80        90       100       110       120

                         130       140       150       160       170       180
a149-1.pep        SPDHAIMVDSALSQQVEILRGPVTLLYSSGNVAGLVDVADGKIPEKMPENGVSGELGLRL
                  |||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
m149-1            SPDHAIMVDTALSQQVEILRGPVTLLYSSGNVAGLVDVADGKIPEKMPENGVSGELGLRL
                         130       140       150       160       170       180

                         190       200       210       220       230       240
a149-1.pep        SSGNLEKLTSGGINIGLGKNFVLHTEGLYRKSGDYAVPRYRNLKRLPDSHADSQTGSIGL
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m149-1            SSGNLEKLTSGGINIGLGKNFVLHTEGLYRKSGDYAVPRYRNLKRLPDSHADSQTGSIGL
                         190       200       210       220       230       240

                         250       260       270       280       290       300
a149-1.pep        SWVGEKGFIGAAYSDRRDQYGLPAHSHEYDDCHADIIWQKSLINKRYLQLYPHLLTEEDI
                  ||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
m149-1            SWVGEKGFIGVAYSDRRDQYGLPAHSHEYDDCHADIIWQKSLINKRYLQLYPHLLTEEDI
                         250       260       270       280       290       300

                         310       320       330       340       350       360
a149-1.pep        DYDNPGLSCGFHDDDDAHAHAHNGKPWIDLRNKRYELRAEWKQPFPGFEALRVHLNRNDY
                  |||||||||||||||:||||:|:|:|||||||||||||||||||||||||||||||||||
m149-1            DYDNPGLSCGFHDDDNAHAHTHSGRPWIDLRNKRYELRAEWKQPFPGFEALRVHLNRNDY
                         310       320       330       340       350       360

                         370       380       390       400       410       420
a149-1.pep        RHDEKAGDAVENFFNNQTQNARIELRHQPIGRLKGSWGVQYLGQKSSALSATSEAVKQPM
                  |||||||||||||||||||||||||||||||||||||||||| |||||||| ||||||||
m149-1            RHDEKAGDAVENFFNNQTQNARIELRHQPIGRLKGSWGVQYLQQKSSALSAISEAVKQPM
                         370       380       390       400       410       420

                         430       440       450       460       470       480
a149-1.pep        LLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASIRYDKALIDRENYYNHPLPDLGAH
                  |||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
m149-1            LLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASIQYDKALIDRENYYNHPLPDLGAH
                         430       440       450       460       470       480

                         490       500       510       520       530       540
a149-1.pep        RQTARSFALSGNWYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKER
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m149-1            RQTARSFALSGNWYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKER
                         490       500       510       520       530       540

                         550       560       570       580       590       600
a149-1.pep        SNNIELALGYEGDRWQYNLALYRNRFGNYIYAQTLNDGRGPKSIEDDSEMKLVRYNQSGA
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m149-1            SNNIELALGYEGDRWQYNLALYRNRFGNYIYAQTLNDGRGPKSIEDDSEMKLVRYNQSGA
                         550       560       570       580       590       600

                         610       620       630       640       650       660
```

```
 451  GGCGACAGTT CCTATCCGAA TTTCTGTCAG GCAGGTAAAG ATTTCGACCG
 501  GCGTTTTGAA GAATTGGGCG CAAAACGGCT GCTCGAACGC GTTGATGCGG
 551  ATTTGGACTT TACCGCCTCC GCAAACGCCT GGACAGATAA TATCGCCGCA
 601  CTCTTAAAAG AAGAAGCCGC AAAAAACCGG GCAACGCCCG CGCCGCAGAC
 651  AACGCCCCCC GCCGGCCTTC AGACGGCACC GGATGGCAGG TACTGCAAGG
 701  CAGCCCCCTT TCCCGCCGCC CTGCTGGCCA ATCAGAAAAT CACCGCCCGC
 751  CAATCCGATA AAGACGTGCG CCACATCGAA ATCGATTTGA GCGGTTCGGA
 801  TTTGCACTAC CTCCCGGGCG ACGCGCTCGG CGTTTGGTTT GACAACGATC
 851  CGGCACTGGT CAGGGAAATC CTAGACCTGC TCGGCATCGA TCCGGCAACG
 901  GAAATACAGG CGGGCGGAAA GATGATGCCG GTTGCGCGCG CACTTTCATC
 951  TCATTTCGAA CTCACGCAAA ACACTCCGGC TTTCGTCAAA GGCTATGCCG
1001  CGTTCGCCCA TTATGAAGAA CTCGATAAAA TCATTGCCGA TAACGCCGTT
1051  TTGCAGGATT TCGTGCAAAA CACGCCTATT GTCGATGTGC TGCACCGCTT
1101  CCCGGCAAGC CTGACGGCAG AACAATTCAT CCGTTTACTG CGTCCGCTTG
1151  CACCCCGTTT GTATTCGATT TCTTCAGCAC AGGCGGAAGT GGGCGATGAA
1201  GTGCATTTAA CTGTCGGCGT GGTTCGTTTT GAACACGAAG GCCGCGCCAG
1251  AACGGGCGGC GCATCGGGTT TCCTTGCCGA CCGGCTGGAA GAGGACGGCA
1301  CGGTGCGCGT GTTTGTGGAA CGCAACGACG GCTTCAGGCT GCCCGAAGAC
1351  AGCCGCAAGC CGATTGTGAT GATCGGCTCG GGCACCGGCG TCGCACCGTT
1401  CCGCGCTTTC GTCCAACAAC GTGCCGCAGA AAATGCGGAA GGCAAAAACT
1451  GGCTGATTTT CGGCAATCCG CATTTTGCCC GTGATTTTCT CTATCAAACC
1501  GAATGGCAGC AGTTTGCCAA AGACGGCTTC CTGCACAGGT ACGATTTCGC
1551  CTGGTCCCGC GATCAGGAAG AAAAAATCTA TGTGCAGGAC AAAATCCGCG
1601  AACAGGCGGA AGGACTTTGG CAATGGCTGC AGGAAGGCGC GCATATCTAT
1651  GTGTGCGGCG ATGCGGCAAA AATGGCAAAA GACGTGGAAG CCGCCTTGCT
1701  GGATGTGATT ATCGGGGCAG GACATTTGGA CGAAGAGGGC GCAGAAGAAT
1751  ATTTGGATAT GCTGCGCGAA GAAAAACGCT ATCAGCGTGA TGTTTATTGA
```

This corresponds to the amino acid sequence <SEQ ID 610; ORF 150>:

```
m150.pep
   1  MQNTNPPLPP LPPEITQLLS GLDAAQWAWL SGYAWAKAGN GASAGLPALQ
  51  TALPAAEPFS VTVLSASQTG NAKSVADKAA DSLEAAGIQV SRAELKDYKA
 101  KNIAGERRLL LVTSTQGEGE PPKEAVVLHK LLNGKKAPKL DKLQFAVLGL
 151  GDSSYPNFCQ AGKDFDRRFE ELGAKRLLER VDADLDFTAS ANAWTDNIAA
 201  LLKEEAAKNR ATPAPQTTPP AGLQTAPDGR YCKAAPFPAA LLANQKITAR
 251  QSDKDVRHIE IDLSGSDLHY LPGDALGVWF DNDPALVREI LDLLGIDPAT
 301  EIQAGGKMMP VARALSSHFE LTQNTPAFVK GYAAFAHYEE LDKIIADNAV
 351  LQDFVQNTPI VDVLHRFPAS LTAEQFIRLL RPLAPRLYSI SSAQAEVGDE
 401  VHLTVGVVRF EHEGRARTGG ASGFLADRLE EDGTVRVFVE RNDGFRLPED
 451  SRKPIVMIGS GTGVAPFRAF VQQRAAENAE GKNWLIFGNP HFARDFLYQT
 501  EWQQFAKDGF LHRYDFAWSR DQEEKIYVQD KIREQAEGLW QWLQEGAHIY
 551  VCGDAAKMAK DVEAALLDVI IGAGHLDEEG AEEYLDMLRE EKRYQRDVY*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 150 shows 91.3% identity over a 369 aa overlap with a predicted ORF (ORF 150.ng) from *N. gonorrhoeae:*

```
m150/g150

                   210       220       230       240       250       260
m150.pep    LLKEEAAKNRATPAPQTTPPAGLQTAPDGRYCKAAPFPAALLANQKITARQSDKDVRHIE
                                          |||| |||||||||||||||||||||||||
g150                                      YCKADPFPAALLANQKITARQSDKDVRHIE
                                                  10        20        30

                   270       280       290       300       310       320
m150.pep    IDLSGSDLHYLPGDALGVWFDNDPALVREILDLLGIDPATEIQAGGKMMPVARALSSHFE
            |||||||||||||||||||||||||| ||||||||:|||||||||| :||| || ||||
g150        IDLSGSDLHYLPGDALGVWFDNDPALVGEILDLLGINPATEIQAGGKTLPVASALLSHFE
                    40        50        60        70        80        90

                   330       340       350       360       370       380
m150.pep    LTQNTPAFVKGYAAFAHYEELDKIIADNAVLQDFVQNTPIVDVLHRFPASLTAEQFIRLL
```

```
1701   GGATGTGATT ATCGGGGCAG GACATTTGGA CGAAGAGGGC GCAGAAGAAT
1751   ATTTGGATAT GCTGCGCGAA GAAAAACGCT ATCAGCGTGA TGTTTATTGA
```

This corresponds to the amino acid sequence <SEQ ID 612; ORF 150.a>:

```
a150.pep
  1   MQNTNPPLPP MPPEITQLLS GLDAAQWAWL SGYAWAKAGN GASAGLPALQ
 51   TALPTAEPFS VTVLSASQTG NAKSVADKAA DSLEAAGIQV SRAELKDYKA
101   KNIAGERRLL LVTSTQGEGE PPEEAVVLHK LLNGKKAPKL DKLQFAVLGL
151   GDSSYPNFCR AGKDFDKRFE ELGAKRLLER VDADLDFAAA ADGWTDNIAA
201   LLKEEAAKNR ATPAPQTTPP AGLQTAPDGR YCKADPFPAA LLANQKITAR
251   QSDKDVRHIE IDLSGSDLHY LPGDALGVWF DNDPALVREI LDLLGIDQAT
301   EIQAGGKTLP VASALLSHFE LTQNTPAFVK GYAPFADDDE LDRIAADNAV
351   LQGFVQSTPI ADVLHRFPAK LTAEQFAGLL RPLAPRLYSI SSSQAEVGDE
401   VHLTVGAVRF EHEGRARAGG ASGFLADRLE EDGTVRVFVE RNDGFRLPED
451   SRKPIVMIGS GTGVAPFRAF VQQRAAENAE GKNWLFFGNP HFARDFLYQT
501   EWQQFAKDGF LHRYDFAWSR DQEEKIYVQD KIREQAEGLW QWLQEGAHIY
551   VCGDAAKMAK DVEAALLDVI IGAGHLDEEG AEEYLDMLRE EKRYQRDVY*
```

m150/a150 94.8% identity in 599 aa overlap

```
                  10        20        30        40        50        60
m150.pep   MQNTNPPLPPLPPEITQLLSGLDAAQWAWLSGYAWAKAGNGASAGLPALQTALPAAEPFS
           ||||||||||:|||||||||||||||||||||||||||||||||||||||||||:|||||
a150       MQNTNPPLPPMPPEITQLLSGLDAAQWAWLSGYAWAKAGNGASAGLPALQTALPTAEPFS
                  10        20        30        40        50        60

                  70        80        90       100       110       120
m150.pep   VTVLSASQTGNAKSVADKAADSLEAAGIQVSRAELKDYKAKNIAGERRLLLVTSTQGEGE
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a150       VTVLSASQTGNAKSVADKAADSLEAAGIQVSRAELKDYKAKNIAGERRLLLVTSTQGEGE
                  70        80        90       100       110       120

                 130       140       150       160       170       180
m150.pep   PPKEAVVLHKLLNGKKAPKLDKLQFAVLGLGDSSYPNFCQAGKDFDRRFEELGAKRLLER
           ||:||||||||||||||||||||||||||||||||||||:||||||:|||||||||||||
a150       PPEEAVVLHKLLNGKKAPKLDKLQFAVLGLGDSSYPNFCRAGKDFDKRFEELGAKRLLER
                 130       140       150       160       170       180

                 190       200       210       220       230       240
m150.pep   VDADLDFTASANAWTDNIAALLKEEAAKNRATPAPQTTPPAGLQTAPDGRYCKAAPFPAA
           |||||||:|:|::||||||||||||||||||||||||||||||||||||||||||| |||||
a150       VDADLDFAAAADGWTDNIAALLKEEAAKNRATPAPQTTPPAGLQTAPDGRYCKADPFPAA
                 190       200       210       220       230       240

                 250       260       270       280       290       300
m150.pep   LLANQKITARQSDKDVRHIEIDLSGSDLHYLPGDALGVWFDNDPALVREILDLLGIDPAT
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
a150       LLANQKITARQSDKDVRHIEIDLSGSDLHYLPGDALGVWFDNDPALVREILDLLGIDQAT
                 250       260       270       280       290       300

                 310       320       330       340       350       360
m150.pep   EIQAGGKMMPVARALSSHFELTQNTPAFVKGYAAFAHYEELDKIIADNAVLQDFVQNTPI
           ||||||| :||| || ||||||||||||||||| ||  :|||:| ||||||| |||:|||
a150       EIQAGGKTLPVASALLSHFELTQNTPAFVKGYAPFADDDELDRIAADNAVLQGFVQSTPI
                 310       320       330       340       350       360

                 370       380       390       400       410       420
m150.pep   VDVLHRFPASLTAEQFIRLLRPLAPRLYSISSAQAEVGDEVHLTVGVVRFEHEGRARTGG
           :||||||||:||||||  ||||||||||||||:|||||||||||||:||||||||||:||
a150       ADVLHRFPAKLTAEQFAGLLRPLAPRLYSISSSQAEVGDEVHLTVGAVRFEHEGRARAGG
                 370       380       390       400       410       420

                 430       440       450       460       470       480
m150.pep   ASGFLADRLEEDGTVRVFVERNDGFRLPEDSRKPIVMIGSGTGVAPFRAFVQQRAAENAE
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a150       ASGFLADRLEEDGTVRVFVERNDGFRLPEDSRKPIVMIGSGTGVAPFRAFVQQRAAENAE
                 430       440       450       460       470       480

                 490       500       510       520       530       540
m150.pep   GKNWLIFGNPHFARDFLYQTEWQQFAKDGFLHRYDFAWSRDQEEKIYVQDKIREQAEGLW
           |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
501  NLEDRGRMFV SPNDKIYEGM IIGIHSRDND LVVNPLKGKK LTNIRASGTD
551  EAVRLTTPIK LTLEGAVEFI DDDELVEITP QSIRLRMRYL SELERRRHFK
601  KLD*
```

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 615>:

```
m151.seq
   1  ATGAAACAAA TCCGCAACAT CGCCATCATC GCCCACGTCG ACCACGGCAA
  51  AACCACATTG GTCGACCAAC TGCTGCGCCA ATCCGGCACA TTCCGCGCCA
 101  ACCAGCAGGT TGACGAGCGC GTGATGGACA GCAACGACCT TGAAAAAGAA
 151  CGCGGCATCA CCATCCTCGC CAAAAACACC GCCATCGATT ACGAAGGCTA
 201  CCACATCAAT ATCGTCGACA CGCCGGGACA CGCCGACTTC GGCGGCGAAG
 251  TAGAGCGCGT TTTGGGGATG GTGGACTGCG TCGTCTTGTT GGTGGACGCG
 301  CAGGAAGGCC CGATGCCGCA AACCCGTTTC GTGACCAAAA AAGCCTTGGC
 351  TTTGGGGCTG AAACCGATTG TCGTCATCAA CAAAATCGAC AAGCCGTCCG
 401  CTCGTCCGAG CTGGGTTATC GACCAAACTT TCGAGCTGTT CGACAACTTG
 451  GGCGCGACCG ACGAGCAGTT GGATTTCCCG ATTGTTTACG CTTCAGGGTT
 501  GAGCGGTTTC GCCAAATTGG AAGAAACCGA CGAGAGCAAC GACATGCGTC
 551  CGCTGTTCGA TACTATCTTA AAATATACGC CTGCACCGAG CGGCAGCGCG
 601  GACGAAACGC TGCAACTGCA AATTTCCCAA CTCGACTACG ACAACTACAC
 651  CGGCCGCCTC GGTATCGGTC GTATCTTGAA CGGACGCATC AAACCCGGCC
 701  AAACCGTTGC CGTCATGAAC CACGATCAGC AAATCGCCCA AGGCCGCATC
 751  AACCAGCTTT TGGGTTTCAA AGGTTTGGAA CGCGTGCCGC TTGAAGAAGC
 801  CGAAGCCGGC GACATCGTGA TTATTTCCGG TATCGAAGAC ATCGGTATCG
 851  GCGTAACCAT CACCGACAAA GACAATCCCA AAGGCCTACC GATGTTGAGC
 901  GTGGACGAAC CGACGCTGAC GATGGACTTT ATGGTCAACA CCAGCCCGCT
 951  GGCGGGTACG GAAGGCAAAT TCGTAACCAG CCGCCAAATC CGCGACCGCC
1001  TGCAAAAAGA ATTGCTGACC AACGTCGCCC TGCGCGTGGA AGATACCGCC
1051  GATGCCGACG TGTTCCGCGT ATCCGGGCGC GGCGAGCTGC ACCTGACCAT
1101  TTTGCTGGAA AACATGCGCC GCGAAGGCTA CGAACTCGCC GTCGGCAAAC
1151  CGCGCGTCGT GTACCGCGAC ATCGACGGTC AAAAATGCGA ACCGTATGAA
1201  AACCTGACCG TGGATGTACC CGACGACAAC CAAGGCGCGG TAATGGAAGA
1251  ACTCGGCCGC CGCCGTGGCG AACTGACTAA TATGGAAAGC GACGGCAACG
1301  GACGCACCCG CCTCGAATAC CATATTCCAG CGCGCGGCTT GATCGGTTTC
1351  CAAGGCGAAT TTATGACCCT GACGCGCGGG GTCGGGCTGA TGAGCCACGT
1401  GTTCGACGAT TACGCGCCCG TCAAACCCGA TATGCCCGGC CGCCACAACG
1451  GCGTGCTGGT GTCCCAAGAG CAGGGCGAGG CAGTCGCTTA CGCCTTGTGG
1501  AATCTGGAAG ACCGCGGCCG TATGTTCGTA TCGCCCAACG ACAAAATCTA
1551  CGAAGGCATG ATTATCGGCA TCCACAGTCG CGACAACGAT TTGGTGGTCA
1601  ACCCGCTCAA AGGCAAAAAA CTTACCAACA TCCGTGCCAG CGGTACCGAC
1651  GAAGCCGTTC GCCTGACCAC GCCAATCAAG CTGACGCTGG AAGGTGCGGT
1701  TGAGTTTATC GACGATGACG AACTCGTTGA AATCACGCCG CAATCCATCC
1751  GTCTGCGCAA GCGTTACTTG AGCGAATTGG AACGCCGCCG CCACTTTAAA
1801  AAGCTGGATT GA
```

This corresponds to the amino acid sequence <SEQ ID 616; ORF 151>:

```
m151.pep
   1  MKQIRNIAII AHVDHGKTTL VDQLLRQSGT FRANQQVDER VMDSNDLEKE
  51  RGITILAKNT AIDYEGYHIN IVDTPGHADF GGEVERVLGM VDCVVLLVDA
 101  QEGPMPQTRF VTKKALALGL KPIVVINKID KPSARPSWVI DQTFELFDNL
 151  GATDEQLDFP IVYASGLSGF AKLEETDESN DMRPLFDTIL KYTPAPSGSA
 201  DETLQLQISQ LDYDNYTGRL GIGRILNGRI KPGQTVAVMN HDQQIAQGRI
 251  NQLLGFKGLE RVPLEEAEAG DIVIISGIED IGIGVTITDK DNPKGLPMLS
 301  VDEPTLTMDF MVNTSPLAGT EGKFVTSRQI RDRLQKELLT NVALRVEDTA
 351  DADVFRVSGR GELHLTILLE NMRREGYELA VGKPRVVYRD IDGQKCEPYE
 401  NLTVDVPDDN QGAVMEELGR RRGELTNMES DGNGRTRLEY HIPARGLIGF
 451  QGEFMTLTRG VGLMSHVFDD YAPVKPDMPG RHNGVLVSQE QGEAVAYALW
 501  NLEDRGRMFV SPNDKIYEGM IIGIHSRDND LVVNPLKGKK LTNIRASGTD
 551  EAVRLTTPIK LTLEGAVEFI DDDELVEITP QSIRLRKRYL SELERRRHFK
 601  KLD*
```

Computer analysis of this amino acid sequence gave the following results:

```
m151.pep      KLDX
              ||||
g151          KLDX
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 617>:

```
a151.seq
   1  ATGAAACAAA TCCGCAACAT CGCCATCATC GCCCACGTCG ACCACGGCAA
  51  AACCACATTG GTCGACCAAC TGCTGCGCCA ATCCGGCACA TTCCGCGCCA
 101  ACCAGCAGGT TGACGAGCGC GTGATGGACA GCAACGACCT TGAAAAAGAA
 151  CGCGGCATCA CCATCCTCGC CAAAAACACC GCCATCGATT ACGAAGGCTA
 201  CCACATCAAT ATCGTCGACA CGCCGGGACA CGCCGACTTC GGCGGCGAAG
 251  TAGAGCGAGT TTTGGGGATG GTGGACTGCG TCGTCTTGTT GGTGGACGCG
 301  CAGGAAGGCC CGATGCCGCA AACCCGTTTC GTGACCAAAA AAGCCTTGGC
 351  TTTGGGGCTG AAACCGATTG TCGTCATCAA TAAAATCGAC AAACCGTCCG
 401  CCCGTCCGAG CTGGGTCATC GACCAAACTT TCGAGCTGTT CGACAACTTG
 451  GGCGCGACCG ACGAGCAGTT GGATTTCCCG ATTGTTTATG CTTCCGGTCT
 501  GTCCGGTTTC GCCAAATTGG AAGAAACCGA CGAGAGCAAC GACATGCGTC
 551  CGCTGTTCGA TACTATCTTA AAATATACGC CTGCACCGAG CGGCAGCGCG
 601  GACGAAACGC TGCAACTGCA AATTTCCCAA CTCGACTACG ACAACTACAC
 651  CGGCCGCCTC GGTATCGGTC GTATCT[illegible] CGGACGTATC AAGCCCGGTC
 701  AAGTTGTTGC CGTCATGAAC CACGATCAAC AAATCGCCCA AGGCCGCATC
 751  AACCAGCTTT TGGGTTTCAA AGGTTTAGAA CGCGTGCCGC TTGAAGAAGC
 801  CGAAGCCGGC GACATCGTGA TTATTTCCGG TATTGAAGAC ATCGGCATCG
 851  GCGTAACCAT CACCGACAAA GACAACCCCA AAGGCCTGCC GATGTTGAGC
 901  GTGGACGAAC CGACGCTGAC GATGGACTTT ATGGTCAACA CCAGCCCGTT
 951  GGCAGGTACG GAAGGCAAAT TCGTAACCAG CCGCCAAATC CGCGACCGCC
1001  TGCAAAAAGA ATTGCTGACC AACGTCGCCC TGCGCGTGGA AGATACCGCC
1051  GATGCCGACG TGTTCCGCGT ATCCGGGCGC GGCGAGCTGC ACCTGACCAT
1101  TTTGCTGGAA AACATGCGCC GCGAAGGCTA CGAACTCGCC GTCGGCAAAC
1151  CGCGCGTCGT GTACCGCGAC ATCGACGGTC AAAAATGCGA ACCGTATGAA
1201  AACCTGACCG TGGACGTACC CGACGACAAC CAAGGCGCGG TAATGGAAGA
1251  ACTCGGCCGC CGCCGTGGCG AACTGACTAA TATGGAAAGC GACGGCAACG
1301  GACGCACCCG CCTCGAATAC CATATTCCAG CGCGCGGCTT GATCGGCTTC
1351  CAAGGCGAAT TTATGACCCT GACGCGCGGG GTCGGGCTGA TGAGCCACGT
1401  GTTCGACGAT TACGCGCCCG TCAAACCCGA TATGCCTGGC CGCCACAACG
1451  GCGTGCTGGT GTCCCAAGAG CAGGGCGAGG CAGTCGCTTA CGCCTTGTGG
1501  AATCTGGAAG ACCGCGGCCG TATGTTCGTA TCGCCCAACG ACAAAATCTA
1551  CGAAGGTATG ATTATCGGCA TCCACAGTCG CGACAACGAT TTGGTGGTCA
1601  ACCCGCTCAA AGGCAAAAAA CTTACCAACA TCCGTGCCAG CGGTACCGAC
1651  GAAGCCGTTC GCCTGACCAC GCCGATTAAG CTGACGCTGG AAGGTGCGGT
1701  CGAGTTTATC GACGATGATG AGCTGGTAGA AATCACGCCG CAATCCATCC
1751  GTCTGCGCAA GCGTTACTTG AGCGAATTGG AACGCCGCCG CCATTTCAAA
1801  AAGCTAGATT GA
```

This corresponds to the amino acid sequence <SEQ ID 618; ORF 151.a>:

```
a151.pep
   1  MKQIRNIAII AHVDHGKTTL VDQLLRQSGT FRANQQVDER VMDSNDLEKE
  51  RGITILAKNT AIDYEGYHIN IVDTPGHADF GGEVERVLGM VDCVVLLVDA
 101  QEGPMPQTRF VTKKALALGL KPIVVINKID KPSARPSWVI DQTFELFDNL
 151  GATDEQLDFP IVYASGLSGF AKLEETDESN DMRPLFDTIL KYTPAPSGSA
 201  DETLQLQISQ LDYDNYTGRL GIGRILNGRI KPGQVVAVMN HDQQIAQGRI
 251  NQLLGFKGLE RVPLEEAEAG DIVIISGIED IGIGVTITDK DNPKGLPMLS
 301  VDEPTLTMDF MVNTSPLAGT EGKFVTSRQI RDRLQKELLT NVALRVEDTA
 351  DADVFRVSGR GELHLTILLE NMRREGYELA VGKPRVVYRD IDGQKCEPYE
 401  NLTVDVPDDN QGAVMEELGR RRGELTNMES DGNGRTRLEY HIPARGLIGF
 451  QGEFMTLTRG VGLMSHVFDD YAPVKPDMPG RHNGVLVSQE QGEAVAYALW
 501  NLEDRGRMFV SPNDKIYEGM IIGIHSRDND LVVNPLKGKK LTNIRASGTD
 551  EAVRLTTPIK LTLEGAVEFI DDDELVEITP QSIRLRKRYL SELERRRHFK
 601  KLD*
```

m151/a151 99.8% identity in 603 aa overlap

```
           10        20        30        40        50        60
```

```
151  CTCGTATTCC GCCTCTGCTG GGGCATTTGG GGCAgcgATA CCGCCCGTTT
201  CTCccgTtTC GTCCGAGGTT GGGCAGGTAT ACGCGGCTAT CTGAAAAAcg
251  gCATTCCCGA ACAtatcCAG CCCGGACACA ACCCCTTGGG CGCACTgatg
301  gtcGTTGCGC TTTTGgccgc cgtcTCATTT CAagtcggcA CGGGGCTTTT
351  Tgccgccaat gaaaacacct tcagcaCCAa cggctacctc aaccatttgg
401  tttccgaaca tacgGGCAGC CTTATACGGA AAATCCACCT CAACTTTTTC
451  AAGCTGCTCG CCGTTTTTTC CGCAGTCCAC ATCGCCGCCG TCGCCGCATA
501  CCGCATATTC AAAAAGAAAA ACCTCGTCCG CCCGATGATA ACCGGCTTCA
551  AATACATCGA AGGCAAAACC TCAATCCGCT TTGCCGGCAA AGCCGCGCTT
601  GCCGCCGCAT TATCGGTTGC CGCGCTTGCC GCAGCCGCCA TCCTGCTCCT
651  GTCCTGA
```

This corresponds to the amino acid sequence <SEQ ID 620; ORF 152.ng>:

```
g152.pep
  1  MKNKTKVWDF PTRLFHWLLA ASLPFMWYSA KAGGDMLQWH TRVGLLVLFL
 51  LVFRLCWGIW GSDTARFSRF VRGWAGIRGY LKNGIPEHIQ PGHNPLGALM
101  VVALLAAVSF QVGTGLFAAN ENTFSTNGYL NHLVSEHTGS LIRKIHLNFF
151  KLLAVFSAVH IAAVAAYRIF KKKNLVRPMI TGFKYIEGKT SIRFAGKAAL
201  AAALSVAALA AAAILLLS*
```

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 621>:

```
m152.seq
  1  ATGAAAAACA AAACCAAAGT CTGGGACCTC CCCACCCGCC TTTTCCACTG
 51  GCTGCTTGCC GCGTCCCTGC CCTTTATGTG GTATAGCGCG AAAGCCGGCG
101  GCGATATGCT GCAATGGCAC ACGCGCGTCG GGCTGTTCGT CCTTTTCCTG
151  CTCGTATTTC GCCTCTGCTG GGGCATTTGG GGCAGCGATA CCGCCCGTTT
201  TTCCCGTTTC GTCCAAGGCT GGGCAGGCAT ACGCGGCTAT CTGAAAAACG
251  GTATTCCCGA ACACATCCAG CCCGGACACA ACCCCTTGGG CGCACTGATG
301  GTCGTTGCGC TTTTGGCCGC CGTGTCCTTC CAAGTCGGCA CCGGGCTTTT
351  TGCCGCCGAT GAAAACACCT TCAGCACCAA CGGCTACCTC AACCATTTGG
401  TTTCCGAACA TACGGGCAGC CTTATGCGGA AAATCCACCT CAACTTTTTC
451  AAGCTGCTCG CCGTTTTTTC TGCAATCCAC ATCGCCGCCG TCGCCGCATA
501  CCGCGTATTC AAAAAGAAAA ACCTCATCCT CCCGATGATA ACCGGCTTCA
551  AATACATCGA AGGCAAAACC TCAATCCGCT TTGCAGGCAA AGCCGCGCTT
601  GCCGCCGCAT TATCGGTTGC CTCGCTTGCC GCAGCCGCCA TCCTGCTCCT
651  GTCCTGA
```

This corresponds to the amino acid sequence <SEQ ID 622; ORF 152>:

```
m152.pep
  1  MKNKTKVWDL PTRLFHWLLA ASLPFMWYSA KAGGDMLQWH TRVGLFVLFL
 51  LVFRLCWGIW GSDTARFSRF VQGWAGIRGY LKNGIPEHIQ PGHNPLGALM
101  VVALLAAVSF QVGTGLFAAD ENTFSTNGYL NHLVSEHTGS LMRKIHLNFF
151  KLLAVFSAIH IAAVAAYRVF KKKNLILPMI TGFKYIEGKT SIRFAGKAAL
201  AAALSVASLA AAAILLLS*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 152 shows 95.4% identity over a 218 aa overlap with a predicted ORF (ORF 152.ng) from *N. gonorrhoeae:*

```
m152/g152

                   10        20        30        40        50        60
m152.pep    MKNKTKVWDLPTRLFHWLLAASLPFMWYSAKAGGDMLQWHTRVGLFVLFLLVFRLCWGIW
            |||||||||:||||||||||||||||||||||||||||||||||||:|||||||||||||
g152        MKNKTKVWDFPTRLFHWLLAASLPFMWYSAKAGGDMLQWHTRVGLLVLFLLVFRLCWGIW
                   10        20        30        40        50        60

                   70        80        90       100       110       120
m152.pep    GSDTARFSRFVQGWAGIRGYLKNGIPEHIQPGHNPLGALMVVALLAAVSFQVGTGLFAAD
            |||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||:
g152        GSDTARFSRFVRGWAGIRGYLKNGIPEHIQPGHNPLGALMVVALLAAVSFQVGTGLFAAN
                   70        80        90       100       110       120
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 625>:

g153.seq

```
   1  atggggtttg cttaCAgtat gacgtatatc gaggtCGGGa taccggaggc
  51  ggcatccgtc ctttCgctGC CCGAGATgat gcgcctgatG GTGTTtCagg
 101  attATGGTTT TttggcCGAA GTGATGTTTG TGctgaCTTT cGGCGcgcCG
 151  GTTCTGTTtC TGCTGCTGTG CCTGTATGTC TATGCCGCGC TGATACGGAA
 201  ACAGGCGTAT CCTGCGCTGC GTTTGGCAAC GCGTGTGATG GTGCGCTTGA
 251  GGCAGGCGAT GATGGTGGAT GTGTTTTTTG TTTCCACTCT GGTGGCGTAT
 301  ATCAAGCTCT CGTCTGTGGC AAAGGTTCGC TTCGGGCCGG CGTTTTATCT
 351  GATGTTCGCG CTGTCGGTTA TGCTGATTCG GACTTCGGTA TCGGTTCCCC
 401  AGCATTGGGT GTATTTCCAA ATCGGGCGGC TGACGGGGAA TAATGCGGTT
 451  CAGACGGCAT CGGAAGGCAA AACCTGTTGC AGCCGCTGCC TGTATTTccg
 501  cgacAGTgcc gaatccCCCT GCGGGGTGTg cgGCGcggaA CTgtacggcg
 551  gacggccgaa aagtCTGAGt atttCgtCGG CGTTTCTgac ggcggcggTT
 601  GTTTTGTATT TCCctgCcaa TATCctgccg attaTGAttt cgtccAATCc
 651  tgccgccacg GAGGcCAACA CCATCTTTAG CGGCATCGCT TATATGTGGG
 701  ACGagggcgA CAGGCTGATT GCGGCGGTTA TTTTCAGCGC GAGTATTTTG
 751  GTGCCGGTGC TGAAGATTGC GGCAATGTCG GTTTTGATTG CGGCGGCACG
 801  GTTCGCTTTG CCGGCGGGCG CAAAGAAATT GTCGCACCTC tacCGCATCA
 851  CCGAAGCGGT CGGCCGCTGG TCGATGATTG ATATTTTTGT GATTATTATT
 901  TTGATGTGTT CGTTCCacaC TTATGCCGCG CGCGTCATTC CGGGCAGTGC
 951  GGCAGTCTAT TTCTGCCTGG TCGTGATTTT GACGATGCTG TCCGCCTATT
1001  ATTTCGACCC GCGCCTGCTT TGGGACAAAC GCGCTTCAGA CGGCATTGCT
1051  TTCAACGAAA CGGAAAAATA TGACTGA
```

This corresponds to the amino acid sequence <SEQ ID 626; ORF 153.ng>:

g153.pep

```
   1  MGFAYSMTYI EVGIPEAASV LSLPEMMRLM VFQDYGFLAE VMFVLTFGAP
  51  VLFLLLCLYV YAALIRKQAY PALRLATRVM VRLRQAMMVD VFFVSTLVAY
 101  IKLSSVAKVR FGPAFYLMFA LSVMLIRTSV SVPQHWVYFQ IGRLTGNNAV
 151  QTASEGKTCC SRCLYFRDSA ESPCGVCGAE LYGGRPKSLS ISSAFLTAAV
 201  VLYFPANILP IMISSNPAAT EANTIFSGIA YMWDEGDRLI AAVIFSASIL
 251  VPVLKIAAMS VLIAAARFAL PAGAKKLSHL YRITEAVGRW SMIDIFVIII
 301  LMCSFHTYAA RVIPGSAAVY FCLVVILTML SAYYFDPRLL WDKRASDGIA
 351  FNETEKYD*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 627>:

m153.seq

```
   1  ATGGCGTTTG CTTACGGTAT GACGTATATC GAGGTCGGGA TACCGGGTGC
  51  GGCATCCGTC CTTTCGCTGC CCGAGATGAT GCGCCTGATG GTGTTTCAGG
 101  ATTATGGTTT TTTGGCCGAA GTGATGTTTG TGCTGACTTT CGGCGCGCCG
 151  GTTCTGTTTC TGCTGCTGTG CCTGTATGTC TATGCCGCGC TGATACGGAA
 201  ACAGGCGTAT CCTGCGCTGC GTTTGGCAAC GCGTGTGATG GTGCGCTTGA
 251  GACAGGCGAT GATGGTGGAT GTGTTTTTTG TTTCCACTTT GGTGGCGTAT
 301  ATCAAGCTCT CGTCTGTGGC AGACGTTCGC TTCGGGCCGG CGTTTTATCT
 351  GATGTTCGCG CTGTCAGTTA TGCTGATTCG GACTTCGGTA TCGGTTCCCC
 401  AGCATTGGGT GTATTTTCAA ATCGGGCGGC TGACGGGGGA TAATGCGGTT
 451  CAGACGGCAT CGGAAGGTAA AACCTGTTGC AGCCGCTGCC TGTATTTCCG
 501  CGACAGTGCC GAATCCCCCT GCGGCGTGTG CGGTGCGGAA CTGTACCGCC
 551  GACGGCCGAA AAGTCTGAGT ATTTCGTCGG CGTTTCTGAC GGCGGCGGTT
 601  ATTTTGTATT TCCCTGCCAA TATCCTGCCG ATTATGATTT CGTCCAATCC
 651  TGCCGCCACG GAGGTCAATA CCATCCTTAA CGGCATCGCT TATATGTGGG
 701  ACGAGGGCGA CAGGCTGATT GCGGCGGTTA TTTTCAGCGC GAGTATTTTG
 751  GTGCCGGTAC TGAAGATTGC GGCAATGTCG GTTTTGATTG CGTCCGCCCG
 801  CTTCGCTTTG CCAACGGGTG CAAAGAAATT GTCGCACCTC TACCGCATCA
 851  CCGAAGCGGT CGGCCGCTGG TCGATGATTG ATATTTTTGT GATTATTATT
 901  TTGATGTGTT CGTTCCACAC TTATGCCGCG CGCGTCATTC CGGGCAGTGC
 951  GGCAGTCTAT TTCTGCCTGG TCGTGATTCT GACGATGCTG TCCGCCTATT
1001  ATTTCGACCC GCGCCTGCTT TGGGACAAAC GCGCTTCAGA CGGCATTGCT
1051  TTCAATGAAA CGGAAAAACA TGACTGA
```

This corresponds to the amino acid sequence <SEQ ID 628; ORF 153>:

m153.pep

```
   1  MAFAYGMTYI EVGIPGAASV LSLPEMMRLM VFQDYGFLAE VMFVLTFGAP
  51  VLFLLLCLYV YAALIRKQAY PALRLATRVM VRLRQAMMVD VFFVSTLVAY
```

```
 851  CCGAAGCGGT CGGCCGCTGG TCGATGATTG ATATTTTTGT GATTATTATT
 901  TTGATGTGTT CGTTCCACAC TTATGCCGCG CGCGTCATTC CGGGCAGTGC
 951  GGCAGTCTAT TTCTGCCTGG TCGTGATTCT GACGATGCTG TCCGCCTATT
1001  ATTTCGACCC GCGCCTGCTT TGGGACAAAC GCGCTTCAGA CGGCATTGCT
1051  TTCAATGAAA CGGAAAAACA TGACTGA
```

This corresponds to the amino acid sequence <SEQ ID 630; ORF 153.a>:

```
a153.pep
      1  MAFAYGMTYI EVGIPGAASV LSLPEMMRLM VFQDYGFLAE VMFVLTFGAP
     51  VLFLLLCLYV YAALIRKQAY PALRLATRVM VRLRQAMMVD VFFVSTLVAY
    101  IKLSSVAEVR FGSAFYLMFA LSVMLIRTSV SVPQHWVYFQ IGRLTGDNAV
    151  QTASEGKTCC SRCLYFRDSA ESPCGVCGAE LYRRRPKSLS ISSAFLTAAV
    201  ILYFPANILP IMISSNPAAT EVNTILNGIA YMWDEGDRLI AAVIFSASIL
    251  VPVLKIAAMS VLIASARFAL PTGAKKLSHL YRITEAVGRW SMIDIFVIII
    301  LMCSFHTYAA RVIPGSAAVY FCLVVILTML SAYYFDPRLL WDKRASDGIA
    351  FNETEKHD*
```

m153/a153 99.7% identity in 358 aa overlap

```
                     10        20        30        40        50        60
m153.pep     MAFAYGMTYIEVGIPGAASVLSLPEMMRLMVFQDYGFLAEVMFVLTFGAPVLFLLLCLYV
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a153         MAFAYGMTYIEVGIPGAASVLSLPEMMRLMVFQDYGFLAEVMFVLTFGAPVLFLLLCLYV
                     10        20        30        40        50        60

                     70        80        90       100       110       120
m153.pep     YAALIRKQAYPALRLATRVMVRLRQAMMVDVFFVSTLVAYIKLSSVAEVRFGPAFYLMFA
             |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
a153         YAALIRKQAYPALRLATRVMVRLRQAMMVDVFFVSTLVAYIKLSSVAEVRFGSAFYLMFA
                     70        80        90       100       110       120

                    130       140       150       160       170       180
m153.pep     LSVMLIRTSVSVPQHWVYFQIGRLTGDNAVQTASEGKTCCSRCLYFRDSAESPCGVCGAE
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a153         LSVMLIRTSVSVPQHWVYFQIGRLTGDNAVQTASEGKTCCSRCLYFRDSAESPCGVCGAE
                    130       140       150       160       170       180

                    190       200       210       220       230       240
m153.pep     LYRRRPKSLSISSAFLTAAVILYFPANILPIMISSNPAATEVNTILNGIAYMWDEGDRLI
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a153         LYRRRPKSLSISSAFLTAAVILYFPANILPIMISSNPAATEVNTILNGIAYMWDEGDRLI
                    190       200       210       220       230       240

                    250       260       270       280       290       300
m153.pep     AAVIFSASILVPVLKIAAMSVLIASARFALPTGAKKLSHLYRITEAVGRWSMIDIFVIII
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a153         AAVIFSASILVPVLKIAAMSVLIASARFALPTGAKKLSHLYRITEAVGRWSMIDIFVIII
                    250       260       270       280       290       300

                    310       320       330       340       350      359
m153.pep     LMCSFHTYAARVIPGSAAVYFCLVVILTMLSAYYFDPRLLWDKRASDGIAFNETEKHDX
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a153         LMCSFHTYAARVIPGSAAVYFCLVVILTMLSAYYFDPRLLWDKRASDGIAFNETEKHDX
                    310       320       330       340       350
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 631>:

```
g154.seq
      1  ATGACTGACA ACAGCCCTCC TCCAAACGGA CACGCTCAAG CACGCGTCCG
     51  CAAAAACAAC accttcctCT CCGCCGTCTG GCTGGTCCCG CTGATCGCGC
    101  TGATTGCCGG CGGCTGGCTT TGGGTTAAGG AAATCCGCAA CAGGGGGCCT
    151  GTGGTTACGC TCTTGATGGA CAGCGCGGAA GGCATCGAAG TCAACAATAC
    201  GGTCATTAAG GTATTGAGCA TCGATGTCGG ACGCGTTACC CGAATCAAAC
    251  TGCGCGACGA CCAAAAAGGC GTGGAAGTTA CTGCCCAACT CAATGCGGAC
    301  GTATCCGGCC TCATCCGCAG CGATACCCAG TTTTGGGTGG TCAAGCCGCG
    351  TATCGACCAA AGCGGcgtAA CCGGTTTGGG TACGCTGCTT TCGGGTTCGT
```

```
1201  TCACCTAAGC TGCGACCGCA TACCGTTTAT GCAGGCGATA CCGTTATCGC
1251  GACCCAGGGC GGCGGTTTGG ACGATTTGCA GGTCAAATTG GCGGATTTGC
1301  TGGACAAGTT CGACAAACTG CCTTTAGATA AGACGGTTGC CGAATTGAAC
1351  GGTTCGCTTG CCGAGCTCAA ATCCACACTC AAATCTGCCA ATGCCGCCCT
1401  AAGCTCCATC GACAAACTGG TCGGCAAACC GCAGACACAA AACATTCCGA
1451  ACGAACTGAA CCAAACCCTG AAAGAGTTGC GCACAACCCT GCAAGGCGTA
1501  TCGCCGCAAT CGCCTATCTA CGGCGACGTA CAAAATACGC TGCAAAGTTT
1551  GGACAAAACT TTAAAAGACG TTCAACCCGT GATTAATACT TTGAAAGAAA
1601  AACCCAACGC GCTGATTTTC AACAGCAGCA GCAAAGACCC TATCCCGAAA
1651  GGAAGCCGAT AA
```

This corresponds to the amino acid sequence <SEQ ID 634 ORF 154.a>:

m154.pep

```
  1  MTDNSPPPNG HAQARVRKNN TFLSAVWLVP LIALIAGGWL WVKEIRNRGP
 51  VVTLLMDSAE GIEVNNTVIK VLSIDVGRVT RIKLRDDQKG VEVTAQLNAD
101  VSGLIRSDTQ FWVVKPRIDQ SGVTGLGTLL SGSYIAFTPG KSDEAKDVFQ
151  VQDIPPVTAI GQSGLRLNLI GKNDRILNVN SPVLYENFMV GQVESAHFDP
201  SDQSVHYTIF IQSPNDKLIH SASRFWLESG INIETTGSGI KLNSAPLPAL
251  LSGAISFDSP KTKNSKNVKS EDSFTLYDSR SEVANLPDDR SLYYTAFFKQ
301  SVRGLTVGSP VEYKGLNVGV VSDVPYFDRN DSLHLFENGW IPVRIRIEPS
351  RLEINADEQS KEHWKQQFQT ALNKGLTATI SSNNLLTGSK MIELNDQPSA
401  SPKLRPHTVY AGDTVIATQG GGLDDLQVKL ADLLDKFDKL PLDKTVAELN
451  GSLAELKSTL KSANAALSSI DKLVGKPQTQ NIPNELNQTL KELRTTLQGV
501  SPQSPIYGDV QNTLQSLDKT LKDVQPVINT LKEKPNALIF NSSSKDPIPK
551  GSR*
```

m154 / g154 97.8% identity in 553 aa overlap

```
                      10        20        30        40        50        60
m154.pep     MTDNSPPPNGHAQARVRKNNTFLSAVWLVPLIALIAGGWLWVKEIRNRGPVVTLLMDSAE
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g154         MTDNSPPPNGHAQARVRKNNTFLSAVWLVPLIALIAGGWLWVKEIRNRGPVVTLLMDSAE
                      10        20        30        40        50        60

                      70        80        90       100       110       120
m154.pep     GIEVNNTVIKVLSIDVGRVTRIKLRDDQKGVEVTAQLNADVSGLIRSDTQFWVVKPRIDQ
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g154         GIEVNNTVIKVLSIDVGRVTRIKLRDDQKGVEVTAQLNADVSGLIRSDTQFWVVKPRIDQ
                      70        80        90       100       110       120

                     130       140       150       160       170       180
m154.pep     SGVTGLGTLLSGSYIAFTPGKSDEAKDVFQVQDIPPVTAIGQSGLRLNLIGKNDRILNVN
             |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
g154         SGVTGLGTLLSGSYIAFTPGKSGEAKDVFQVQDIPPVTAIGQSGLRLNLIGKNDRILNVN
                     130       140       150       160       170       180

                     190       200       210       220       230       240
m154.pep     SPVLYENFMVGQVESAHFDPSDQSVHYTIFIQSPNDKLIHSASRFWLESGINIETTGSGI
             ||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
g154         SPVLYENFMVGQIESAHFDPSDQSVHYTIFIQSPNDKLIHSASRFWLESGINIETTGSGI
                     190       200       210       220       230       240

                     250       260       270       280       290       300
m154.pep     KLNSAPLPALLSGAISFDSPKTKNSKNVKSEDSFTLYDSRSEVANLPDDRSLYYTAFFKQ
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
g154         KLNSAPLPALLSGAISFDSPKTKNSKNVKSEDSFTLYDSRSEIANLPDDRSLYYTAFFKQ
                     250       260       270       280       290       300

                     310       320       330       340       350       360
m154.pep     SVRGLTVGSPVEYKGLNVGVVSDVPYFDRNDSLHLFENGWIPVRIRIEPSRLEINADEQS
             |||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
g154         SVRGLTVGSPVEYKGLNVGMVSDVPYFDRNDSLHLFENGWIPVRIRIEPSRLEINADEQS
                     310       320       330       340       350       360

                     370       380       390       400       410       420
m154.pep     KEHWKQQFQTALNKGLTATISSNNLLTGSKMIELNDQPSASPKLRPHTVYAGDTVIATQG
             ||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||:|
g154         KEHWKQQFQTALNKGLTATISSNNLLTGGKMIELNDQPSASPKLRPHTVYAGDTVIATRG
```

```
151  VQDIPPVTAI GQSGLRLNLI GKNDRILNVN SPVLYENFMV GQVESAHFDP
201  SDQSVHYTIF IQSPNDKLIH SASRFWLESG INIETTGSGI KLNSAPLPAL
251  LSGAISFDSP KTKNSKNVKS EDSFTLYDSR SEVANLPDDR SLYYTAFFKQ
301  SVRGLTVGSP VEYKGLNVGV VSDVPYFDRN DSLHLFENGW IPVRIRIEPS
351  RLEINADEQS KEHWKQQFQT ALNKGLTATI SSNNLLTGSK MIELNDQPSA
401  SPKLRPHTVY AGDTVIATQG GGLDDLQVKL ADLLDKFDKL PLDKTVAELN
451  GSLAELKSTL KSANAALSSI DKLVGKPQTQ NIPNELNQTL KELRTTLQGV
501  SPQSPIYGDV QNTLQSLDKT LKDVQPVINT LKEKPNALIF NSSSKDPIPK
551  GSR*
```

m154/a154   100.0% identity in 553 aa overlap

```
                   10        20        30        40        50        60
m154.pep   MTDNSPPPNGHAQARVRKNNTFLSAVWLVPLIALIAGGWLWVKEIRNRGPVVTLLMDSAE
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a154       MTDNSPPPNGHAQARVRKNNTFLSAVWLVPLIALIAGGWLWVKEIRNRGPVVTLLMDSAE
                   10        20        30        40        50        60

                   70        80        90       100       110       120
m154.pep   GIEVNNTVIKVLSIDVGRVTRIKLRDDQKGVEVTAQLNADVSGLIRSDTQFWVVKPRIDQ
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a154       GIEVNNTVIKVLSIDVGRVTRIKLRDDQKGVEVTAQLNADVSGLIRSDTQFWVVKPRIDQ
                   70        80        90       100       110       120

                  130       140       150       160       170       180
m154.pep   SGVTGLGTLLSGSYIAFTPGKSDEAKDVFQVQDIPPVTAIGQSGLRLNLIGKNDRILNVN
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a154       SGVTGLGTLLSGSYIAFTPGKSDEAKDVFQVQDIPPVTAIGQSGLRLNLIGKNDRILNVN
                  130       140       150       160       170       180

                  190       200       210       220       230       240
m154.pep   SPVLYENFMVGQVESAHFDPSDQSVHYTIFIQSPNDKLIHSASRFWLESGINIETTGSGI
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a154       SPVLYENFMVGQVESAHFDPSDQSVHYTIFIQSPNDKLIHSASRFWLESGINIETTGSGI
                  190       200       210       220       230       240

                  250       260       270       280       290       300
m154.pep   KLNSAPLPALLSGAISFDSPKTKNSKNVKSEDSFTLYDSRSEVANLPDDRSLYYTAFFKQ
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a154       KLNSAPLPALLSGAISFDSPKTKNSKNVKSEDSFTLYDSRSEVANLPDDRSLYYTAFFKQ
                  250       260       270       280       290       300

                  310       320       330       340       350       360
m154.pep   SVRGLTVGSPVEYKGLNVGVVSDVPYFDRNDSLHLFENGWIPVRIRIEPSRLEINADEQS
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a154       SVRGLTVGSPVEYKGLNVGVVSDVPYFDRNDSLHLFENGWIPVRIRIEPSRLEINADEQS
                  310       320       330       340       350       360

                  370       380       390       400       410       420
m154.pep   KEHWKQQFQTALNKGLTATISSNNLLTGSKMIELNDQPSASPKLRPHTVYAGDTVIATQG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a154       KEHWKQQFQTALNKGLTATISSNNLLTGSKMIELNDQPSASPKLRPHTVYAGDTVIATQG
                  370       380       390       400       410       420

                  430       440       450       460       470       480
m154.pep   GGLDDLQVKLADLLDKFDKLPLDKTVAELNGSLAELKSTLKSANAALSSIDKLVGKPQTQ
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a154       GGLDDLQVKLADLLDKFDKLPLDKTVAELNGSLAELKSTLKSANAALSSIDKLVGKPQTQ
                  430       440       450       460       470       480

                  490       500       510       520       530       540
m154.pep   NIPNELNQTLKELRTTLQGVSPQSPIYGDVQNTLQSLDKTLKDVQPVINTLKEKPNALIF
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a154       NIPNELNQTLKELRTTLQGVSPQSPIYGDVQNTLQSLDKTLKDVQPVINTLKEKPNALIF
                  490       500       510       520       530       540
```

```
 301  TTGGTCGAAG CCTTGCGCGC CAAGAAAGTG AACGCGCTGG CGATGGATAT
 351  GGTGCCCCGC ATTTCGCGCG CGCAGGCTTT GGACGCTTTG TCTTCGATGG
 401  CAAACATCAG CGGCTACCGC GCCGTAATTG AAGCCGCCAA CGCCTTCGGC
 451  CGTTTCTTCA CCGGTCAAAT TACCGCCGCC GGCAAAGTGC CGCCCGCGCA
 501  GGTTTTGGTG ATTGGTGCAG GTGTGGCAGG TTTGGCGGCG ATCGGTACGG
 551  CAAACTCGCT CGGCGCAGTG GTACGCGCGT TCGATACCCG CTTGGAAGTG
 601  GCGGAACAAA TCGAATCGAT GGGCGGCAAG TTCCTGAAAC TCGACTTCCC
 651  ACAAGAATCG GGCGGCAGCG GAGACGGCTA CGCCAAAGTG ATGAGCGACG
 701  AATTTATCGC AGCCGAGATG AAGCTCTTTG CCGAGCAGGC GAAAGAAGTG
 751  GACATCATCA TCACCACCGC CGCCATTCCG GGCAAACCCG CGCCCAAGCT
 801  GATTACCAAA GAAATGGTGG AAAGCATGAA ATCCGGCTCC GTCATCGTCG
 851  ATTTGGCGGC GGCGACGGGC GGCAACTGCG AACTCACCCG CCCGGGCGAA
 901  TTGTCCGTAA CCGGCAACGG CGTGAAAATC ATCGGCTACA CCGACATGGC
 951  AAACCGCCTT GCCGGACAGT CTTCCCAGCT TTACGCCACC AACTTGGTCA
1001  ACCTGACCAA GCTGTTAAGC CCGAACAAAG ACGGCGAAAT CACGTTGGAC
1051  TTCGAAGACG TGATTATCCG CAACATGACC GTTACCCACG ACGGCGAAAT
1101  CACCTTCCCG CCTCCGCCGA TTCAAGTTTC CGCCCAGCCG CAGCAAACGC
1151  CGTCTGAAAA AGCCGTGCCT GCCGCCAAGC CCGAGCCAAA ACCCGTTCCC
1201  CTGTGGAAAA AACTCGCGCC CGCCGTCATC GCCGCCGTCT TGGTACTGTG
1251  GGTCGGCGCG GTCGCACCCG CAGCATTCCT GAACCACTTT ATCGTGTTCG
1301  TTCTCGCCTG CGTCATCGGC TACTACGTCG TCTGGAACGT CAGCCACTCG
1351  CTGCACACAC CGCTGATGTC GGTAACCAAC GCCATCTCCG GCATCATCGT
1401  CGTCGGCGCG CTGCTGCAAA TCGGTCAGGG CAACGGCTTC GTTTCGCTGC
1451  TGTCGTTTGT TGCCATCCTG ATTGCCGGCA TCAACATCTT CGGCGGCTTT
1501  GCGGTAACAC GGCGTATGCT GAATATGTTT AAGAAAGGGT AA
```

This corresponds to the amino acid sequence <SEQ ID 640; ORF 155>:

```
m155.pep
  1  MKIGIPRESL SGETRVACTP ATVALLGKLG FETVVESGAG LAASLDDAAY
 51  QTAGATVADK AAVWVCPLIY KVNAPSEQEL PLLNEGQTIV SFLWPRQNEA
101  LVEALRAKKV NALAMDMVPR ISRAQALDAL SSMANISGYR AVIEAANAFG
151  RFFTGQITAA GKVPPAQVLV IGAGVAGLAA IGTANSLGAV VRAFDTRLEV
201  AEQIESMGGK FLKLDFPQES GGSGDGYAKV MSDEFIAAEM KLFAEQAKEV
251  DIIITTAAIP GKPAPKLITK EMVESMKSGS VIVDLAAATG GNCELTRPGE
301  LSVTGNGVKI IGYTDMANRL AGQSSQLYAT NLVNLTKLLS PNKDGEITLD
351  FEDVIIRNMT VTHDGEITFP PPPIQVSAQP QQTPSEKAVP AAKPEPKPVP
401  LWKKLAPAVI AAVLVLWVGA VAPAAFLNHF IVFVLACVIG YYVVWNVSHS
451  LHTPLMSVTN AISGIIVVGA LLQIGQGNGF VSLLSFVAIL IAGINIFGGF
501  AVTRRMLNMF KKG*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 155 shows 97.9% identity over a 513 aa overlap with a predicted ORF (ORF 155.ng) from *N. gonorrhoeae:*

```
m155 / g155  97.9% identity in 513 aa overlap
                      10        20        30        40        50        60
m155.pep     MKIGIPRESLSGETRVACTPATVALLGKLGFETVVESGAGLAASLDDAAYQTAGATVADK
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g155         MKIGIPRESLSGETRVACTPATVALLGKLGFETVVESGAGLAASLDDAAYQTAGATVADK
                      10        20        30        40        50        60

                      70        80        90       100       110       120
m155.pep     AAVWVCPLIYKVNAPSEQELPLLNEGQTIVSFLWPRQNEALVEALRAKKVNALAMDMVPR
             ||||:||||||||||| ||||:||||||||||||||||||||||||||||||||||||||
g155         AAVWACPLIYKVNAPSEGELPLLKEGQTIVSFLWPRQNEALVEALRAKKVNALAMDMVPR
                      70        80        90       100       110       120

                     130       140       150       160       170       180
m155.pep     ISRAQALDALSSMANISGYRAVIEAANAFGRFFTGQITAAGKVPPAQVLVIGAGVAGLAA
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g155         ISRAQALDALSSMANISGYRAVIEAANAFGRFFTGQITAAGKVPPAQVLVIGAGVAGLAA
```

```
1201  CTGTGGAAAA AACTCGCGCC CGCCNTNATC GCCGCCGTGT TGGTACTGTG
1251  GGTCGGCGCG GTCGCACCCG CAGCATTCCT GAACCACTTT ATCGTCTTCG
1301  TCCTCGCCTG CGTCATCGGC TACTATGTCG TTTGGAACGT CAGCCACTCG
1351  CTGCACACAC CGCTGATGTC GGTGACCAAC GCCATTTCCG GCATCATCGT
1401  CGTCGGCGCG CTGCTGCAAA TCGGTCAGGG CAACGGCTTC GTTTCGCTGC
1451  TGTCGTTTGT TGCCATCCTG ATTGCCAGCA TCAACATCTT CGGCGGCTTC
1501  TTTGTAACGC GGCGGATGCT GAATATGTTT AGGAAAGGGT AA
```

This corresponds to the amino acid sequence <SEQ ID 642; ORF 155.a>:

```
a155.pep
     1  MKIGIPRESL SGETRVACTP ATVALLGKLG FETVVESGAG LAASLDDAAY
    51  QAAGATVADK AAVWAYPLIY KVNAPSEDEL PLLKEGQTIV SFLWPRQNEA
   101  LVEALRAKKV NALAMDMVPR ISRAQALDXL SXMANISGYR AVIEAANAFG
   151  RXFTGQITAA GKVPPAQVLV IGAGVAGLAA IGTANSLGAV VRVFDTRLXV
   201  AEQLESMGGK FLKLDFPQES GGSGDGYAKV MSDEFIAAEM KLFAEQAKEV
   251  DIIITTAAIP GKPAPKXXXK EMVESMKPGS VIVDLAAATG GNCELTKQGE
   301  LFVTGNGVKI IGYTDMANRL AGQSSQLYAT NLVNLTKLLS PNKDGEITLD
   351  FEDVIIRNMT VTRDGEITFP PPPIQVSAQP QQTPSEKAAP AAKPEPKPVP
   401  LWKKLAPAXI AAVLVLWVGA VAPAAFLNHF IVFVLACVIG YYVVWNVSHS
   451  LHTPLMSVTN AISGIIVVGA LLQIGQGNGF VSLLSFVAIL IASINIFGGF
   501  FVTRRMLNMF RKG*
```

m155/a155 95.3% identity in 513 aa overlap

```
                         10        20        30        40        50        60
m155.pep         MKIGIPRESLSGETRVACTPATVALLGKLGFETVVESGAGLAASLDDAAYQTAGATVADK
                 |||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
a155             MKIGIPRESLSGETRVACTPATVALLGKLGFETVVESGAGLAASLDDAAYQAAGATVADK
                         10        20        30        40        50        60

                         70        80        90       100       110       120
m155.pep         AAVWVCPLIYKVNAPSEQELPLLNEGQTIVSFLWPRQNEALVEALRAKKVNALAMDMVPR
                 ||||: |||||||||||:|||||:||||||||||||||||||||||||||||||||||||
a155             AAVWAYPLIYKVNAPSEDELPLLKEGQTIVSFLWPRQNEALVEALRAKKVNALAMDMVPR
                         70        80        90       100       110       120

                        130       140       150       160       170       180
m155.pep         ISRAQALDALSSMANISGYRAVIEAANAFGRFFTGQITAAGKVPPAQVLVIGAGVAGLAA
                 |||||||| || ||||||||||||||||||| ||||||||||||||||||||||||||||
a155             ISRAQALDXLSXMANISGYRAVIEAANAFGRXFTGQITAAGKVPPAQVLVIGAGVAGLAA
                        130       140       150       160       170       180

                        190       200       210       220       230       240
m155.pep         IGTANSLGAVVRAFDTRLEVAEQIESMGGKFLKLDFPQESGGSGDGYAKVMSDEFIAAEM
                 ||||||||||||:||||| ||||:||||||||||||||||||||||||||||||||||||
a155             IGTANSLGAVVRVFDTRLXVAEQLESMGGKFLKLDFPQESGGSGDGYAKVMSDEFIAAEM
                        190       200       210       220       230       240

                        250       260       270       280       290       300
m155.pep         KLFAEQAKEVDIIITTAAIPGKPAPKLITKEMVESMKSGSVIVDLAAATGGNCELTRPGE
                 ||||||||||||||||||||||||||  :|||||||| ||||||||||||||||||: ||
a155             KLFAEQAKEVDIIITTAAIPGKPAPKXXXKEMVESMKPGSVIVDLAAATGGNCELTKQGE
                        250       260       270       280       290       300

                        310       320       330       340       350       360
m155.pep         LSVTGNGVKIIGYTDMANRLAGQSSQLYATNLVNLTKLLSPNKDGEITLDFEDVIIRNMT
                 | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a155             LFVTGNGVKIIGYTDMANRLAGQSSQLYATNLVNLTKLLSPNKDGEITLDFEDVIIRNMT
                        310       320       330       340       350       360

                        370       380       390       400       410       420
m155.pep         VTHDGEITFPPPPIQVSAQPQQTPSEKAVPAAKPEPKPVPLWKKLAPAVIAAVLVLWVGA
                 ||:|||||||||||||||||||||||||:||||||||||||||||||| |||||||||||
a155             VTRDGEITFPPPPIQVSAQPQQTPSEKAAPAAKPEPKPVPLWKKLAPAXIAAVLVLWVGA
                        370       380       390       400       410       420
```

```
g156         FAPFAAAVLTAHATGNAGQATVNTLAGLFILFRLAFIWCYIADKAALRSLMWAGGFACTV
                     70        80        90       100       110       120

m156.pep     GLFVAAAX
             ||||||||
g156         GLFVAAAX
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 647>:

```
a156.seq
      1  ATGACTTTCG CCTATTGGTG TATTCTGATT GCCTACCTAT TGCCGCTTTT
     51  TTGTGCGGCG TATGCCAAAA AAGCGGGCGG ATTCCGGTTT AAAGACAACC
    101  ACAATCCGCG CGATTTTCTG GCGCGCACGC AAGGCACAGC CGCCCGTGCC
    151  CACGCCGCGC AGCAAAACGG TTTTGAAGCC TTTGCACCGT TTGCAGCCGC
    201  CGTTTTGACG GCACACGCAA CCGGCAATGC CGGACAAGCA ACCGTCAACA
    251  CGCTTGCCGG CCTGTTCATC CTGTTCCGCC TCGCCTTTAT CTGGTGCTAC
    301  ATCGCAGACA AAGCAGCATT ACGCTCGCTG ATGTGGGTGG GCGGATTTGT
    351  CTGCACCGTC GGGCTGTTTG TCGTGGCTGC TTGA
```

This corresponds to the amino acid sequence <SEQ ID 648; ORF 156.a>:

```
a156.pep
      1  MTFAYWCILI AYLLPLFCAA YAKKAGGFRF KDNHNPRDFL ARTQGTAARA
     51  HAAQQNGFEA FAPFAAAVLT AHATGNAGQA TVNTLAGLFI LFRLAFIWCY
    101  IADKAALRSL MWVGGFVCTV GLFVVAA*

m156/a156    90.6% identity in 127 aa overlap

                     10        20        30        40        50        60
m156.pep     MTFAYWCILIACLLPLFCAAYAKKAGGFRFKDNHNPRGFLAHTQGAAARAHAAQQNGFEA
             |||||||||| ||||||||||||||||||||||||||| |||:|||:|||||||||||||
a156         MTFAYWCILIAYLLPLFCAAYAKKAGGFRFKDNHNPRDFLARTQGTAARAHAAQQNGFEA
                     10        20        30        40        50        60

                     70        80        90       100       110       120
m156.pep     FAPFAAAVLTAHATGNAAQSTINTLACLFILFRLAFIWCYIADKAAMRSLMWAGGFACTV
             ||||||||||||||||:|:|:|||| |||||||||||||||||||:|||||:|||:|||
a156         FAPFAAAVLTAHATGNAGQATVNTLAGLFILFRLAFIWCYIADKAALRSLMWVGGFVCTV
                     70        80        90       100       110       120

m156.pep     GLFVAAAX
             ||||:|||
a156         GLFVVAAX
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 649>:

```
g157.seq
      1  atgaggaacg aggAAAAACg cgccctgcgc cgcgaattgC gCgGgcggcg
     51  ttcgcAAATg GGgcgagacg tGCGggCGGC GGCGgCgatA Aaaatcaacc
    101  gcctgctcaa aCGTtatatc AAGCGCggtc gGaAaatcgG CGTGTATTgg
    151  cCGATGGGCA AGGAATTGcg TTTGGGCGgc tTtgtcCGCG CGGCGCAAAA
    201  ACGCgGCGCA AAactctatc tgccttATAT CGAACCGCAC ACGCGGCGGA
    251  TGTGGTTTAC GCCGTATCCT GAACGCGGAA TGGAACGGGA ACGCAAGCGC
    301  GGTAGGGCGA AGCTGCATGT CCCTCAGTTT GCAGGGCGCA AAATCCGCGT
    351  GCACGGTTTG TCGGTATTGC TCGTCCCGCT TGTCGGCATA GACCGCGAAG
    401  GCTACCGTTT GGGGCAGGCA GGCGGCTATT ACGATGCGAC GCTTTCGGCG
    451  ATGAAATACC GTTTGCAGGC GAAAACCGTG GGCGTGGGCT TTGCCTGCCA
    501  GTTGGTGGAC AGGCTCCCAC GCGAGGCGCA CGACCTGCCG CTGGACGGTT
    551  TTGTATCGGA AGCGGGGATA TTGTGTTTTT AG
```

```
101  GCCTGCTCAA ACGTTATATC AAGCGTGGTC GGAAAATCGG CGTGTATTGG
151  CCGATGGGCA AGGAATTGCG TTTGGACGGC TTTGTCCGCG CGGCGCAAAA
201  ACGCGGTGCA AAACTTTATC TGCCTTATAT CGAACCGCGT TCGCGGCGGA
251  TGTGGTTTAC GCCGTATCCT GAAAGCGGAA TGGAACGGGA GCGCATACGG
301  GGCAGGGCGA AGTTGAACGT GCCGCAGTTT GCAGGGCGCA AAATCCGCGT
351  GCACGGTTTG TCGGTATTGC TCGTCCCGCT TGTCGGCATA GACCGCGAGG
401  GCTACCGCTT AGGACAGGCA GGCGGCTATT ACGATGCGAC GCTTGCGGCG
451  ATGAAATACC GTTTGCAGGC AAAAACCGTG GGCGTGGGCT TTGCCTGCCA
501  GTTTGTGGAC AGGCTGCCGC GCGAACCGCA CGATCTGCTG CTGGACGGTT
551  TTGTGTCGGA GGCGGGGATA TTGTGCTTTT AG
```

This corresponds to the amino acid sequence <SEQ ID 654; ORF 157.a>:

a157.pep

```
  1  MRNEEKHALR RELRRARAQM GHQGRLAAGQ TINRLLKRYI KRGRKIGVYW
 51  PMGKELRLDG FVRAAQKRGA KLYLPYIEPR SRRMWFTPYP ESGMERERIR
101  GRAKLNVPQF AGRKIRVHGL SVLLVPLVGI DREGYRLGQA GGYYDATLAA
151  MKYRLQAKTV GVGFACQFVD RLPREPHDLL LDGFVSEAGI LCF*
```

m157/a157 82.4% identity in 193 aa overlap

```
                          10        20        30        40        50        60
m157.pep          MRNEEKRALRRELRGRRSQMGRDVRAAATVKINHLLKRYIKKGRKIGVYWPMGKELRLDG
                  ||||||:|||||||  |:|||:: | ||   ||:|||||||:||||||||||||||||||
a157              MRNEEKHALRRELRRARAQMGHQGRLAAGQTINRLLKRYIKRGRKIGVYWPMGKELRLDG
                          10        20        30        40        50        60

                          70        80        90       100       110       120
m157.pep          FVRAAQKRGAELYLPYIEPRSRRMWFTPYPADGVKQERKRGRAKLHVPQFAGRKKRVHDL
                  ||||||||||:||||||||||||||||||| :|:::|| ||||||:|||||||| ||| |
a157              FVRAAQKRGAKLYLPYIEPRSRRMWFTPYPESGMERERIRGRAKLNVPQFAGRKIRVHGL
                          70        80        90       100       110       120

                         130       140       150       160       170       180
m157.pep          NLLLVPVVGMDRLGYRLGQAGGYYDATLSAMKYRLQAKTVGVGFACQLVDRLPVEAHDRS
                  ::||||:||:|| |||||||||||||||:||||||||||||||||||:||||| | ||
a157              SVLLVPLVGIDREGYRLGQAGGYYDATLAAMKYRLQAKTVGVGFACQFVDRLPREPHDLL
                         130       140       150       160       170       180

                         190
m157.pep          LDGFVSEAGILCFX
                  ||||||||||||||
a157              LDGFVSEAGILCFX
                         190
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 655>:

g158.seq

```
  1  ATGAAAACCA ATTCAGAAGA ACTGACCGTA TTTGTTCAAG TGGTGGAAAG
 51  CGGCAGCTTC AGCCGTGCGG CGgagcAGTT GGAGAtggCA AATTCTGCCG
101  TAAGCCGCAT CGTCAAACGG CTGGAGGAAA AGTTGGGCGT GAAcCTGCtc
151  aACCGCACCA CGCGGCAACT CAATCTGACG GAAGAAGGCG CGCAATATTT
201  CCGCCGCGCG CAGAGAATCC TGCAAGAAAT GGCAGCGGCG GAAACCGAAA
251  TGCTGGCAGT GCACGAAGTA CCGCAAGGCG TGTTGCGCGT GGATTCCGCG
301  ATGCCgatgg TGCTGCATCT GCTGGCGCCG CTGGCAGCAA AATTCAACGA
351  ACGCTATCCG CATATCCgaC TTTCGCTCGT TTCTTCCGAa ggctatatca
401  atctGattGA Acgcaaagtc gAtatTGCCT TACGGGCCGG AGAATTGGAC
451  GATTCCGGGC TGCGTGCACG CCATCTGTTT GACAGCCACT TCCGCGtagt
501  cgCCAGTCCT GAATATTTAG CAAAACACGG CACGCCACAA TCTGCAGAAG
551  atcTTGCCAA CCATCAATGT TTAGGCTTCA CAGAACCCGG TTCTCTAAAT
601  ACATGGGCGG TTTTAGAtgC GCAGGGAAAT CCCTATAAAA TTTCACCGCA
651  CTTTACCGCC AGCAGCGGTG AAATCTTACG CTCGTTGTGC CTTTCAAGtt
701  gCGGTATTGC TTGCTTATCA GATTTTTTGG TTGACAACGA CATCACTGAA
751  GGAAAGTTAA TTCCcctatt cgCCGAACAA ACCTCCAATA AAACACACCC
```

```
                        130       140       150       160       170       180
m158.pep    HIRLSLVSSEGYINLIERKVDIALRAGELDDSGLRARHLFDSRFRVIASPEYLAKHGTPQ
            ||||||||||||||||||||||||||||||||||||||||||:|||:|||||||||||||
g158        HIRLSLVSSEGYINLIERKVDIALRAGELDDSGLRARHLFDSHFRVVASPEYLAKHGTPQ
                        130       140       150       160       170       180

                        190       200       210       220       230       240
m158.pep    STEELAGHQCLGFTEPGSLNTWAVLDAQGNPYKISPHFTASSGEILRSLCLSGCGIVCLS
            |:|:||:||||||||||||||||||||||||||||||||||||||||||||:||||:|||
g158        SAEDLANHQCLGFTEPGSLNTWAVLDAQGNPYKISPHFTASSGEILRSLCLSSCGIACLS
                        190       200       210       220       230       240

                        250       260       270       280       290       300
m158.pep    DFLVDNDIAEGKLIPLLAEQTSDKTHPFNAVYYSDKAVNLRLRVFLDFLVEELGNNLCGX
            ||||||||:||||||||:||||:|||||||||||||||||||||||||||:|||:|:
g158        DFLVDNDITEGKLIPLFAEQTSNKTHPFNAVYYSDKAVNLRLRVFLDFLVKELGKNMNRT
                        250       260       270       280       290       300

g158        NTKX
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 659>:

a158.seq

```
  1  ATGAAAACCA ATTCAGAAGA ACTGACCGTA TTTGTTCAAG TGGTGGAAAG
 51  CGGCAGCTTC AGCCGTGCGG CGGAGCAGTT GGCGATGGCA AATTCTGCCG
101  TAAGCCGCAT CGTCAAACGG CTGGAGGAAA AGTTGGGTGT GAACCTGCTC
151  AACCGCACCA CGCGGCAACT CAGTCTGACG GAAGAAGGCG CGCAATATTT
201  CCGCCGCGCG CAGAGAATCC TGCAAGAAAT GGCAGCGGCG GAAACCGAAA
251  TGCTGGCAGT GCACGAAATA CCGCAAGGCG TGTTGCGCGT GGATTCCGCG
301  ATGCCGATGG TGCTGCATCT GCTGGCGCCG CTGGCAGCAA AATTCAACGA
351  ACGCTATCCG CATATCCGAC TTTCGCTCGT TTCTTCCGAA GGCTATATCA
401  ATCTGATTGA ACGCAAAGTC GATATTGCCT TACGGGCCGG AGAATTGGAC
451  GATTCCGGGC TGCGTGCACG CCATCTGTTT GACAGCCGCT TCCGCGTAAT
501  CGCCAGTCCT GAATACCTGG CAAAACACGG CACGCCGCAA TCTACAGAAG
551  AGCTTGCCGG CCACCAATGT TTAGGCTTCA CCGAACCCGG TTCTCTAAAT
601  ACATGGGCGG TTTTAGATGC GCAGGGAAAT CCCTATAAGA TTTCACCGCA
651  CTTTACCGCC AGCAGCGGTG AAATCTTACG CTCGTTGTGC CTTTCAGGTT
701  GCGGTATTGC TTGCTTATCA GATTTTTTGG TTGACAACGA CATCGCTGAA
751  GGAAAGTTAA TTCCCCTGCT CGCCGAACAA ACCTCCAATA AAACGCACCC
801  CTTTAATGCT GTTTATTACA GCGATAAAGC CGTCAACCTC CGCTTACGCG
851  TATTTTTGGA TTTTTTAGTG GAGGAACTGG GAAACAATCT CTGTGGATAA
```

This corresponds to the amino acid sequence <SEQ ID 660; ORF 158.a>:

a158.pep

```
  1  MKTNSEELTV FVQVVESGSF SRAAEQLAMA NSAVSRIVKR LEEKLGVNLL
 51  NRTTRQLSLT EEGAQYFRRA QRILQEMAAA ETEMLAVHEI PQGVLRVDSA
101  MPMVLHLLAP LAAKFNERYP HIRLSLVSSE GYINLIERKV DIALRAGELD
151  DSGLRARHLF DSRFRVIASP EYLAKHGTPQ STEELAGHQC LGFTEPGSLN
201  TWAVLDAQGN PYKISPHFTA SSGEILRSLC LSGCGIACLS DFLVDNDIAE
251  GKLIPLLAEQ TSNKTHPFNA VYYSDKAVNL RLRVFLDFLV EELGNNLCG*
```

m158/a158   99.0% identity in 299 aa overlap

```
                        10        20        30        40        50        60
m158.pep    MKTNSEELTVFVQVVESGSFSRAAEQLAMANSAVSRIVKRLEEKLGVNLLNRTTRQLSLT
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a158        MKTNSEELTVFVQVVESGSFSRAAEQLAMANSAVSRIVKRLEEKLGVNLLNRTTRQLSLT
                        10        20        30        40        50        60

                        70        80        90       100       110       120
m158.pep    EEGAQYFRRAQRILQEMAAAETEMLAVHEIPQGVLSVDSAMPMVLHLLAPLAAKFNERYP
            ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
a158        EEGAQYFRRAQRILQEMAAAETEMLAVHEIPQGVLRVDSAMPMVLHLLAPLAAKFNERYP
                        70        80        90       100       110       120
```

```
501  CGTCCTGCTG GTGCTTATCC TGCGCGCCTA TCTCGAACAG GATAAGGATG
551  TCGAACTCTC GGGCGTATTG AAAGGTTGGC AGGACAAACG TTTGGGACAT
601  TTAATCCAAA AGGTGATAGA CAAACCGGAA GACGAATGGA ATGTCGACAA
651  AATGGTGGCG GCTGCCAATA TGTCGCGCGC GCAACTGATG CGCCGTTTCA
701  AAAGCCGGGT CGGACTCAGC CCGCACGCCT TTGTGAACCA TATCCGCCTG
751  CAAAAAGGCG CGTTGCTGCT GAAAAAAAAC CCGGATTCGG TTTTGTCGGT
801  CGCACTGTCG GTAGGCTTTC AGTCGGAAAC GCACTTCGGC AAGGCGTTCA
851  AACGGCAATA TCACGTTTCG CCGGGTCAAT ACCGGAAAGA AggCGGGCAA
901  AAATAA
```

This corresponds to the amino acid sequence <SEQ ID 664; ORF 160>:

```
m160.pep
  1  MDILDKLVDF AQLTGSVDVQ CLLGGQWSVR HETLQREGLV HIVTSGSGYL
 51  CIDGETSPRP VSTGDIVFFP RGLGHVLSHD GKCGESLQPD MRQHGAFTVK
101  QCGNGQDMSL FCARFRYDTH ADLMNGLPET VFLNIAHPSL QYVVSMLQLE
151  SKKPLTGTVS MVNALSSVLL VLILRAYLEQ DKDVELSGVL KGWQDKRLGH
201  LIQKVIDKPE DEWNVDKMVA AANMSRAQLM RRFKSRVGLS PHAFVNHIRL
251  QKGALLLKKN PDSVLSVALS VGFQSETHFG KAFKRQYHVS PGQYRKEGGQ
301  K*
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

```
m160 / g160  93.4% identity in 301 aa overlap

                     10        20        30        40        50        60
m160.pep     MDILDKLVDFAQLTGSVDVQCLLGGQWSVRHETLQREGLVHIVTSGSGYLCIDGETSPRP
             ||||||||:||||||:|||||||||||   |||||||||||||:||||||||||||||||
g160         MDILDKLVDLAQLTGSADVQCLLGGQW---HETLQREGLVHIVTAGSGYLCIDGETSPRP
                     10        20           30        40        50

                     70        80        90       100       110       120
m160.pep     VSTGDIVFFPRGLGHVLSHDGKCGESLQPDMRQHGAFTVKQCGNGQDMSLFCARFRYDTH
             |:|||||||||||||||||||| ||||||:||:|:| ||||||| |||||||||||||||
g160         VGTGDIVFFPRGLGHVLSHDGKYGESLQPDIRQNGTFMVKQCGNGLDMSLFCARFRYDTH
              60        70        80        90       100       110

                    130       140       150       160       170       180
m160.pep     ADLMNGLPETVFLNIAHPSLQYVVSMLQLESKKPLTGTVSMVNALSSVLLVLILRAYLEQ
             ||||||||||||||||||||||||||||||:||||||||:|||| ||||||||||||||
g160         ADLMNGLPETVFLNIAHPSLQYVVSMLQLESEKPLTGTVSVVNALPSVLLVLILRAYLEQ
             120       130       140       150       160       170

                    190       200       210       220       230       240
m160.pep     DKDVELSGVLKGWQDKRLGHLIQKVIDKPEDEWNVDKMVAAANMSRAQLMRRFKSRVGLS
             |||||||||||||||||||||||||||||||||:||||||||||||||||||||:||||
g160         DKDVELSGVLKGWQDKRLGHLIQKVIDKPEDEWNIDKMVAAANMSRAQLMRRFKSQVGLS
             180       190       200       210       220       230

                    250       260       270       280       290       300
m160.pep     PHAFVNHIRLQKGALLLKKNPDSVLSVALSVGFQSETHFGKAFKRQYHVSPGQYRKEGGQ
             ||||||||||||||||||:||||| |||||||||||||||||||||||||||||||||||
g160         PHAFVNHIRLQKGALLLKKTPDSVLEVALSVGFQSETHFGKAFKRQYHVSPGQYRKEGGQ
             240       250       260       270       280       290

m160.pep     KX
             ||
g160         KX
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 665>:

```
a160.seq
  1  ATGGACATTC TGGACAAACT GGTCGATTTC GCCCAATTGA CGGGCAGTGT
 51  GGATGTGCAG TGCCTTTTGG GCGGACAATG GTCGGTACGG CATGAAACCT
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 667>:

```
g161.seq
   1  ATGGATACCG CAAAAAAAGA CATTTTAGGA TCGGGCTGGA TGCTGGTGGC
  51  GGCGGCCTGC TTCACCGTTA TGAACGTATT GATTAAAGAG GCATCGGCAA
 101  AATTTGCCCT CGGCAGCGGC GAATTGGTCT TTTGGCGCAT GCTGTTTTCA
 151  ACCGTTACGC TCGGTGCTGC CGCCGTATTG CGGCGCGACA CCTTCCGCAC
 201  GCCCCATTGG AAAAACCACT TAAACCGCAG TATGGTCGGG ACGGGGGCGA
 251  TGCTGCTGCT GTTTTACGCG GTAACGCATC TGCCTTTGAC AACCGGCGTT
 301  ACCCTGAGTT ACACCTCGTC GATTTTTttg GCGGTATTTT CCTTCCTGAT
 351  TTTGAAAGAA CGGATTTCCG TTTACACGCA GGCGGTGCTG CTCCTTGGTT
 401  TTGCCGGCGT GGTATTGCTG CTTAATCCCT CGTTCCGCAG CGGTCAGGAA
 451  CCGGCGGCAC TCGCCGGGCT GGCGGGCGGC GCGATGTCCG GCTGGGCGTA
 501  TTTGAAAGTG CGCGAACTGT CTTTGGCGGG CGAACCCGGC TGGCGCGTCG
 551  TGTTTTACCT TTCCGCAACC GGCGTGGCGA TGTCGTCggt ttgggcgacg
 601  Ctgaccggct ggCACAcccT GTCCTTTcca tcggcagttt ATCtgtCGGG
 651  CATCGGCGTG tccgcgCtgA TTGCCCAaCT GtcgatgAcg cGCGcctaca
 701  aaGTCGGCGA CAAATTCACG GTTGCCTCGC tttcctaTAt gaccgtcGTC
 751  TTTTCCGCCC TGTCTGCCGC ATTTTTTCTg ggcgaagagc ttttctggCA
 801  GGAAATACTC GGTATGTGCA TCATTATcct CAGCGGCATT TTGAGCAGCA
 851  TCCGCCCCAT TGCCTTCAAA CAGCGGCTGC AAGCCCTCTT CCGCCAAAGA
 901  TAA
```

This corresponds to the amino acid sequence <SEQ ID 668; ORF 161.ng>:

```
g161.pep
   1  MDTAKKDILG SGWMLVAAAC FTVMNVLIKE ASAKFALGSG ELVFWRMLFS
  51  TVTLGAAAVL RRDTFRTPHW KNHLNRSMVG TGAMLLLFYA VTHLPLTTGV
 101  TLSYTSSIFL AVFSFLILKE RISVYTQAVL LLGFAGVVLL LNPSFRSGQE
 151  PAALAGLAGG AMSGWAYLKV RELSLAGEPG WRVVFYLSAT GVAMSSVWAT
 201  LTGWHTLSFP SAVYLSGIGV SALIAQLSMT RAYKVGDKFT VASLSYMTVV
 251  FSALSAAFFL GEELFWQEIL GMCIIILSGI LSSIRPIAFK QRLQALFRQR
 301  *
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 669>:

```
m161.seq
   1  ATGGATACCG CAAAAAAAGA CATTTTAGGA TCGGGCTGGA TGCTGGTGGC
  51  GGCGGCCTGC TTTACCATTA TGAACGTATT GATTAAAGAG GCATCGGCAA
 101  AATTTGCCCT CGGCAGCGGC GAATTGGTCT TTTGGCGCAT GCTGTTTTCA
 151  ACCGTTGCGC TCGGGGCTGC CGCCGTATTG CGTCGGGACA mCTTCCGCAC
 201  GCCCCATTGG AAAAACCACT TAAACCGCAG TATGGTCGGG ACGGGGGCGA
 251  TGCTGCTGCT GTTTTACGCG GTAACGCATC TGCCTTTGGC CACTGGCGTT
 301  ACCCTGAGTT ACACCTCGTC GATTTTTTTG GCGGTATTTT CCTTCCTGAT
 351  TTTGAAAGAA CGGATTTCCG TTTACACGCA GGCGGTGCTG CTCCTTGGTT
 401  TTGCCGGCGT GGTATTGCTG CTTAATCCCT CGTTCCGCAG CGGTCAGGAA
 451  ACGGCGGCAC TCGCCGGGCT GGCGGGCGGC GCGATGTCCG GCTGGGCGTA
 501  TTTGAAAGTG CGCGAACTGT CTTTGGCGGG CGAACCCGGC TGGCGCGTCG
 551  TGTTTTACCT TTCCGTGACA GGTGTGGCGA TGTCGTCGGT TTGGGCGACG
 601  CTGACCGGCT GGCACACCCT GTCCTTTCCA TCGGCAGTTT ATCTGTCGTG
 651  CATCGGCGTG TCCGCGCTGA TTGCCCAACT GTCGATGACG CGCGCCTACA
 701  AAGTCGGCGA CAAATTCACG GTTGCCTCGC TTTCCTATAT GACCGTCGTT
 751  TTTTCCGCTC TGTCTGCCGC ATTTTTTCTG GGCGAAGAGC TTTTCTGGCA
 801  GGAAATACTC GGTATGTGCA TCATCATCCT CAGCGGTATT TTGAGCAGCA
 851  TCCGCCCCAC TGCCTTCAAA CAGCGGCTGC AATCCCTGTT CCGCCAAAGA
 901  TAA
```

This corresponds to the amino acid sequence <SEQ ID 670; ORF 161>:

```
m161.pep
   1  MDTAKKDILG SGWMLVAAAC FTIMNVLIKE ASAKFALGSG ELVFWRMLFS
  51  TVALGAAAVL RRDXFRTPHW KNHLNRSMVG TGAMLLLFYA VTHLPLATGV
 101  TLSYTSSIFL AVFSFLILKE RISVYTQAVL LLGFAGVVLL LNPSFRSGQE
```

```
851  TCCGCCCCAC TGCCTTCAAA CAGCGGCTGC AATCCCTGTT CCGCCAAAGA
901  TAA
```

This corresponds to the amino acid sequence <SEQ ID 672; ORF 161.a>:

```
a161.pep
  1  MDTAKKDILG SGWMLVAAAC FTIMNVLIKE ASAKFALGSG ELVFWRMLFS
 51  TVALGAAAVL RRDTFRTPHW KNHLNRSMVG TGAMLLLFYA VTHLPLATGV
101  TLSYTSSIFL AVFSFLILKE RISVYTQAVL LLGFAGVVLL LNPSFRSGQE
151  TAALAGLAGG AMSGWAYLKV RELSLAGEPG WRVVFYLSVT GVAMSSVWAT
201  LTGWHTLSFP SAVYLSCIGV SALIAQLSMT RAYKVGDKFT VASLSYMTVV
251  FSALSAAFFL AEELFWQEIL GMCIIILSGI LSSIRPTAFK QRLQSLFRQR
301  *
```

m161/a161 99.3% identity in 300 aa overlap

```
                  10        20        30        40        50        60
m161.pep  MDTAKKDILGSGWMLVAAACFTIMNVLIKEASAKFALGSGELVFWRMLFSTVALGAAAVL
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a161      MDTAKKDILGSGWMLVAAACFTIMNVLIKEASAKFALGSGELVFWRMLFSTVALGAAAVL
                  10        20        30        40        50        60

                  70        80        90       100       110       120
m161.pep  RRDXFRTPHWKNHLNRSMVGTGAMLLLFYAVTHLPLATGVTLSYTSSIFLAVFSFLILKE
          |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a161      RRDTFRTPHWKNHLNRSMVGTGAMLLLFYAVTHLPLATGVTLSYTSSIFLAVFSFLILKE
                  70        80        90       100       110       120

                 130       140       150       160       170       180
m161.pep  RISVYTQAVLLLGFAGVVLLLNPSFRSGQETAALAGLAGGAMSGWAYLKVRELSLAGEPG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a161      RISVYTQAVLLLGFAGVVLLLNPSFRSGQETAALAGLAGGAMSGWAYLKVRELSLAGEPG
                 130       140       150       160       170       180

                 190       200       210       220       230       240
m161.pep  WRVVFYLSVTGVAMSSVWATLTGWHTLSFPSAVYLSCIGVSALIAQLSMTRAYKVGDKFT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a161      WRVVFYLSVTGVAMSSVWATLTGWHTLSFPSAVYLSCIGVSALIAQLSMTRAYKVGDKFT
                 190       200       210       220       230       240

                 250       260       270       280       290       300
m161.pep  VASLSYMTVVFSALSAAFFLGEELFWQEILGMCIIILSGILSSIRPTAFKQRLQSLFRQR
          ||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
a161      VASLSYMTVVFSALSAAFFLAEELFWQEILGMCIIILSGILSSIRPTAFKQRLQSLFRQR
                 250       260       270       280       290       300

m161.pep  X
          |
a161      X
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 673>:

```
g163.seq
  1  ATGGTTATTT TGACGACTTT GTTTTTTGTG TGTGTTTTGG TGGTATTGGT
 51  TTTAACCGTG CCGGATCAGG TGCAGATGTG gctCGACCGG GCAAAAGAAG
101  TCATTTTTAC CGAGTTCAGC TGGTTTTATG TTTTAACGTT TTCCATTTTt
151  ctgGGTTTtc tgctGATACT CTCGGTCAGC GGTTTGGGAA ACATcagGCT
201  AGGACGGGAT GAAGATGTGC CGGAATTCGG CTTCCTGTCG TGGCTGGCGA
251  TGCTGTTTGC GGCCGGGATG GGCGTGGGCC TGATGTTTTT CGGCGTGGCA
301  GAGCCGTTGA TGCATTATTT TTCGGACATT ACGGTCGGCG CGCCGGAACA
351  CAGGCAGCAG CAGGCATTGC TGCACACGGT GTTCCATTGG GGCGTTCACG
401  CCTGGTCGGT GTACGGTACG ATTGCATTGG CTTTGGCTTA TTTCGGTTTC
```

```
 651  CGTGCAGGTT TTGATTATCG CCGCCGTCAT GTCCCTCGCC GTCGTTTCGG
 701  CAATATCCGG CGTGGGGAAG GGCGTGAAGG TGTTGAGCGA GTTGAACCTG
 751  GGCCTTGCGT TTTTGCTGCT GTTTTTTGTT TTGGCGGCGG GACCCACTGT
 801  TTACCTGTTG TCGGCATTCG GCGACAACAT AGGGAACTAC CTCGGAAATC
 851  TGGTGCGCCT CAGTTTTAAA ACTTATGCGT ACGAACGGGA ACACAAGCCG
 901  TGGTTTGAAT CTTGGACGGT GCTTTATTGG GCGTGGTGGT GTTCTTGGGC
 951  GCCGTTTGTG GGTTTGTTTA TCGCGCGCAT TTCAAAGGGG CGCACCATCC
1001  GCGAGTTTGT CTTCGGGGTT TTGCTCATCC CCGGCCTGTT CGGCGTTTTG
1051  TGGTTTACCG TCTTCGGCAA TACGGCGATT TGGCTGAATG ACGGGGTTGC
1101  GGGGGGAATG CTCGAAAAGA TGACCTCCTC TCCGGAAACG CTGCTTTTTA
1151  AATTCTTTAA TTACCTCCCC CTGCCCGAAT TGACGAGCAT CGTCAGCCTG
1201  CTGGTCATTT CTCTGTTTTT TGTAACTTCT GCCGATTCCG GGATTTATGT
1251  CCTGAACAAT ATTACCTCTC GGGACAAAGG CTTGAGCGCG CCACGGTGGC
1301  AGGCGGTTAT GTGGGGCGTG CTGATGTCTG CCGTTGCCGT TTTGCTGATG
1351  CGCTCGGGCG GACTCGGCAA CCTGCAGTCT ATGACCCTGA TTGTTTCCCT
1401  GCCGTTTGCC CTGCTGATGC TGATAATGTG TTTCAGCCTG TGGAAAGGCT
1451  TGAGTGCGGA TAAGAAATAT TTTGAGACCC GGGTTAACCC TACCAGTGTA
1501  TTTTGGACGG GCGGCAAGTG GAAAGAACGG CTGGTGCAGA TAATGAGCCA
1551  GACGCAGGAG CAGGATATTT TAAAATTCCT CAAACAGACT GCATCGCCCG
1601  CTATGCACGA GTTGCAACGG GAGCTTTCGG AAGAATACGG CTTGAGCGTC
1651  CGGGTCGATA AAATGTTTCA TCGGGACGAG CCCGCAATCG AGTTCGTCAT
1701  TCGGAAAGAG ACGATGCGCG ATTTTATGTA CGGGATTAAG TCTGTCGGGC
1751  AGGATGTATC CGACCAGTTG ATTAACGACG GCAAGCTGCC GCATATCCGG
1801  CATCAGACAA CTTACAAACC CTACGCTTAT TTTTTCGACG GGCGCGTCGG
1851  GTACGATGTG CAGTATATGA ACAAGGACGA GCTGATTGCC GACATTTTGA
1901  AAAACTACGA ACGTTATTTG ATGTTGTTGG ATGATGTCGG TCAGGAACTG
1951  ATGGCGCACG AGCAGGTGGA ATTGGCAGAG TAA
```

This corresponds to the amino acid sequence <SEQ ID 676; ORF 163>:

```
m163.pep
   1  MVILTTLFFV CVLVVLVLTV PDQVQMWLDR AKEVIFTEFS WFYVLTFSIF
  51  LGFLLILSVS SLGNIRLGRD EDVPEFGFLS WLAMLFAAGM GVGLMFFGVA
 101  EPLMHYFSDI TAGTPEHRQQ QALLHTVFHW GVHAWSVYGT IALALAYFGF
 151  RYKLPLALRS CFYPLLKEKI SGRFGDAIDI MALLATFFGI ITTLGFGASQ
 201  LGAGLQEMGW IAENSFSVQV LIIAAVMSLA VVSAISGVGK GVKVLSELNL
 251  GLAFLLLFFV LAAGPTVYLL SAFGDNIGNY LGNLVRLSFK TYAYEREHKP
 301  WFESWTVLYW AWWCSWAPFV GLFIARISKG RTIREFVFGV LLIPGLFGVL
 351  WFTVFGNTAI WLNDGVAGGM LEKMTSSPET LLFKFFNYLP LPELTSIVSL
 401  LVISLFFVTS ADSGIYVLNN ITSRDKGLSA PRWQAVMWGV LMSAVAVLLM
 451  RSGGLGNLQS MTLIVSLPFA LLMLIMCFSL WKGLSADKKY FETRVNPTSV
 501  FWTGGKWKER LVQIMSQTQE QDILKFLKQT ASPAMHELQR ELSEEYGLSV
 551  RVDKMFHRDE PAIEFVIRKE TMRDFMYGIK SVGQDVSDQL INDGKLPHIR
 601  HQTTYKPYAY FFDGRVGYDV QYMNKDELIA DILKNYERYL MLLDDVGQEL
 651  MAHEQVELAE *
```

Computer analysis of this amino acid sequence gave the following results:
<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

```
m163 / g163   98.6% identity in 660 aa overlap

                      10        20        30        40        50        60
m163.pep      MVILTTLFFVCVLVVLVLTVPDQVQMWLDRAKEVIFTEFSWFYVLTFSIFLGFLLILSVS
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g163          MVILTTLFFVCVLVVLVLTVPDQVQMWLDRAKEVIFTEFSWFYVLTFSIFLGFLLILSVS
                      10        20        30        40        50        60

                      70        80        90       100       110       120
m163.pep      SLGNIRLGRDEDVPEFGFLSWLAMLFAAGMGVGLMFFGVAEPLMHYFSDITAGTPEHRQQ
              :|||||||||||||||||||||||||||||||||||||||||||||||||||:|:|||||
g163          GLGNIRLGRDEDVPEFGFLSWLAMLFAAGMGVGLMFFGVAEPLMHYFSDITVGAPEHRQQ
                      70        80        90       100       110       120

                     130       140       150       160       170       180
```

```
 201  CGGACGGGAT GAAGATGTGC CGGAATTCGG CTTCCTGTCG TGGCTGGCGA
 251  TGCTGTTTGC GGCCGGGATG GGCGTGGGTC TGATGTTTTT CGGCGTGGCA
 301  GAGCCGTTGA TGCATTATTT TTCGGACATT ACGGCCGGCA CGCCGGAACA
 351  CAGGCAGCAG CAGGCATTGC TGCACACGGT GTTCCATTGG GGCGTTCACG
 401  CTTGGTCGGT GTACGGTACG ATTGCATTGG CTTTGGCTTA TTTCGGTTTC
 451  CGCTACAAGC TGCCGCTTGC CCTGCGTTCT TGTTTTTACC CCCTGTTGAA
 501  AGAAAAAATT TCCGGAAGGT TCGGCGATGC CATTGATATT ATGGCGTTGC
 551  TTGCTACTTT TTTCGGCATC ATCACCACAT TGGGGTTCGG GGCTTCGCAA
 601  CTGGGCGCCG GATTGCAGGA AATAGGCTGG ATTGCCGAAA ACAGCTTCAG
 651  CGTGCAGGTT TTGATTATCG CCGCCGTCAT GTCCCTCGCC GTCGTTTCGG
 701  CAATATCCGG CGTGGGGAAG GGTGTGAAGG TGTTGAGCGA GTTGAACCTG
 751  GGTCTTGCGT TTTTGCTGCT GTTTTTTGTT TTGGCGGCGG GTCCCACTGT
 801  TTACCTGTTG TCGGCATTCG GCGACAACAT AGGGAACTAC CTCGGAAATC
 851  TGGTGCGCCT CAGTTTTAAA ACTTATGCGT ACGAACGGGA ACACAAGCCG
 901  TGGTTTGAAT CTTGGACGGT GCTTTATTGG GCGTGGTGGT GTTCTTGGGC
 951  GCCGTTTGTG GGTTTGTTTA TCGCGCGCAT TTCAAAGGGG CGCACCATCC
1001  GCGAGTTTGT CTTCGGGGTT TTGCTCATCC CCGGCCTGTT CGGCGTTTTG
1051  TGGTTTACCG TCTTCGGCAA TACGGCGATT TGGCTGAATG ACGGGGTTGC
1101  GGGGGGAGTG CTCGAAAAGA TGACCTCCTC TCCGGAAACG CTGCTTTTTA
1151  AATTCTTTAA TTACCTCCCC CTGCCCGAAT TGACGAGCAT CGTCAGCCTG
1201  CTGGTCATTT CTCTGTTTTT TGTAACTTCT GCCGATTCCG GGATTTATGT
1251  CCTGAACAAT ATTACCTCTC GGGACAAAGG CTTGAGCGCG CCACGGTGGC
1301  AGGCGGTTAT GTGGGGCGTG CTGATGTCTG CCGTTGCCGT TTTGCTGATG
1351  CGCTCGGGCG GACTCGGCAA CCTGCAGTCT ATGACCCTGA TTGTTTCCCT
1401  GCCGTTTGCC CTGCTGATGC TGATAATGTG TTTCAGCCTG TGGAAAGGAT
1451  TGAGTGCGGA TAAGAAATAT TTTGAGACCC GGGTTAACCC TACCAGTGTA
1501  TTTTGGACGG GCGGCAAGTG GAAAGAACGG CTGGTGCAGA TAATGAGCCA
1551  GACGCAGGAG CAGGATATTT TAAAATTCCT CAAACATACC GCATCGCCCG
1601  CTATGCACGA GTTACAACGG GAGCTTTCGG AAGAATACGG CTTGAGCGTC
1651  CGGGTCGATA AGATGTTTCA TCAGGACGAG CCCGCAATCG AGTTCGTCAT
1701  TCGGAAAGAG ACGATGCGCG ATTTTATGTA CGGGATTAAG TCTGTCGGGC
1751  AGGATGTATC CGACCAGTTG ATTAACGACG GCAAGCTGCC GCATATCCGG
1801  CATCAGACAA CTTACAAACC CTACGCTTAT TTTTTCGACG GGCGCGTCGG
1851  GTACGATGTG CAGTATATGA ACAAGGACGA GCTGATTGCC GACATTTTGA
1901  AAAACTACGA ACGTTATTTG ATGTTGTTGG ATGATGTCGG TCAGGAACTG
1951  ATGGCGCACG AGCAGGTGGA ATTGGCAGAG TAA
```

This corresponds to the amino acid sequence <SEQ ID 678; ORF 163.a>:

```
a163.pep
      1  MVILTTLFFV CVLVVLVLTV PDQVQMWLDR AKEVIFTEFS WFYVLTFSIF
     51  LGFLLILSVS SLGNIRLGRD EDVPEFGFLS WLAMLFAAGM GVGLMFFGVA
    101  EPLMHYFSDI TAGTPEHRQQ QALLHTVFHW GVHAWSVYGT IALALAYFGF
    151  RYKLPLALRS CFYPLLKEKI SGRFGDAIDI MALLATFFGI ITTLGFGASQ
    201  LGAGLQEIGW IAENSFSVQV LIIAAVMSLA VVSAISGVGK GVKVLSELNL
    251  GLAFLLLFFV LAAGPTVYLL SAFGDNIGNY LGNLVRLSFK TYAYEREHKP
    301  WFESWTVLYW AWWCSWAPFV GLFIARISKG RTIREFVFGV LLIPGLFGVL
    351  WFTVFGNTAI WLNDGVAGGV LEKMTSSPET LLFKFFNYLP LPELTSIVSL
    401  LVISLFFVTS ADSGIYVLNN ITSRDKGLSA PRWQAVMWGV LMSAVAVLLM
    451  RSGGLGNLQS MTLIVSLPFA LLMLIMCFSL WKGLSADKKY FETRVNPTSV
    501  FWTGGKWKER LVQIMSQTQE QDILKFLKHT ASPAMHELQR ELSEEYGLSV
    551  RVDKMFHQDE PAIEFVIRKE TMRDFMYGIK SVGQDVSDQL INDGKLPHIR
    601  HQTTYKPYAY FFDGRVGYDV QYMNKDELIA DILKNYERYL MLLDDVGQEL
    651  MAHEQVELAE *
```

m163/a163 99.4% identity in 660 aa overlap

```
                      10        20        30        40        50        60
m163.pep     MVILTTLFFVCVLVVLVLTVPDQVQMWLDRAKEVIFTEFSWFYVLTFSIFLGFLLILSVS
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a163         MVILTTLFFVCVLVVLVLTVPDQVQMWLDRAKEVIFTEFSWFYVLTFSIFLGFLLILSVS
                      10        20        30        40        50        60

                      70        80        90       100       110       120
m163.pep     SLGNIRLGRDEDVPEFGFLSWLAMLFAAGMGVGLMFFGVAEPLMHYFSDITAGTPEHRQQ
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 679>:

```
g164.seq  (partial)
   1  ..ATGAACACAT TTTTGAAAAA CAGCGAATAC GCGTATATCC TGAACGACTG
  51    CAAGGCGCGC TTCCTGTTCG CCTCGGCCGG CCTGTCAAAA GAATTGGCGG
 101    GCCTGAAGGC GCAAACGCCC GTCGAAAAAA TCATTTGGAC GGACAAAAGC
 151    CGGCCGGCCG GCGAAACGGC GGAAGGCGAT GCCTTTTTTG AAAACGTGCG
 201    CCGCTTCCCC GAAAAACCCG ACTTGGGCCG CCAACCCCGG ATAAATGATT
 251    TGGCACACAT CATCTACACC TCCGGCACGA CGGGGCATCC CAAAGGCGCG
 301    CTAATCAGTT ACGCCAACCT GTTCGCCAAC CTGAACGGCA TCGAACGCAT
 351    CTTtaaAATT TCCAAACGCG ACCGCTTTAT CGTTTTCctg ccgatgTTCC
 401    ACAGCTTCAC GCTGACGGCT ATGGTGCTGC TGCCGATTTA TATGGCGTGT
 451    TCGATTATTT TGGTCAAAtc cgttttCCCc ttttccaacG TTTTGAAACA
 501    GGCCCTGCTC AAACGCGCAA CCGTGTTTTT GGGCGTACCC GCGATTTACA
 551    CCGCGATGAG CAAGGCAAAA ATCCCTTGGT ATTTCAGATG GTTCAACCGC
 601    ATCCGCCTGT TTATCAGCGG CGGCGCGCCT TTGGCGGAAC AAACCATCCT
 651    CGATTTTAAA GCCAAGTTCC CCCGCGCCAA ATTGCTGGAA GGCTACGGAC
 701    TGAGCGAAGC CTCGCCCGTC GTCGCCGTCA 'TACGCCCGA ACGGCAAAAA
 751    GCCCGCAGCG TCGGCATCCC CCTGCCCGGT TTGGAAGCCA AAGCCGTCGA
 801    TGAAGAATTG GTCGAAGTGC CGCGCGGCGA AGTGGGCGAA CTGATCGTCA
 851    GGGGCGGTTC GGTGATGCGG GGCTACCTCA ATATGCCTGC CGCCACCGAT
 901    GAAACCATCG TCAACGGCTG GTTGAAAACG GGCGATTTCG TTACCATAGA
 951    CGAGGACGGC TTTATCTTTA TCGTCGACCG CAAAAAAGAT TTGATTATTT
1001    CCAAAGGTCA AAACGTCTAT CCGCGCGAGA TCGAAGAAGA AATCCACAAA
1051    CTCGATGCCG TCGAAGCCGC CGCCGTCATC GGCGTGAAAG ACCGTTATGC
1101    CGACGAGGAA ATCGTCGCCT TCGTCCAATT GAAGGAAGGT ATGGATTTGG
1151    GCGAGGACGA aatccgccgc caccTGCGTA CCGTGCTGGC AAATTTCAAA
1201    ATCCCCAAAC AGATCCACTT TAAAGACGGG CTGCCGCGCA ACGCTACGGG
1251    CAAAGTATTG AAACGGGTGC TGAAGGAGCA GTTTGAAGGA AACAAATGA
```

This corresponds to the amino acid sequence <SEQ ID 680; ORF 164.ng>:

```
g164.pep  (partial)
   1  ..MNTFLKNSEY AYILNDCKAR FLFASAGLSK ELAGLKAQTP VEKIIWTDKS
  51    RPAGETAEGD AFFENVRRFP EKPDLGRQPR INDLAHIIYT SGTTGHPKGA
 101    LISYANLFAN LNGIERIFKI SKRDRFIVFL PMFHSFTLTA MVLLPIYMAC
 151    SIILVKSVFP FSNVLKQALL KRATVFLGVP AIYTAMSKAK IPWYFRWFNR
 201    IRLFISGGAP LAEQTILDFK AKFPRAKLLE GYGLSEASPV VAVNTPERQK
 251    ARSVGIPLPG LEAKAVDEEL VEVPRGEVGE LIVRGGSVMR GYLNMPAATD
 301    ETIVNGWLKT GDFVTIDEDG FIFIVDRKKD LIISKGQNVY PREIEEEIHK
 351    LDAVEAAAVI GVKDRYADEE IVAFVQLKEG MDLGEDEIRR HLRTVLANFK
 401    IPKQIHFKDG LPRNATGKVL KRVLKEQFEG NK*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 681>:

```
m164.seq
   1  ATGAACCGGA CTTATGCCAA TTTCTACGAA ATGCTCGCCG CCGCCTGCCG
  51  CAAAAACGGA AACGGCACGG CAGTGTTCGA CGGCAAGGAA AAAACCGCCT
 101  ACCGCGCGCT CAAGCAGGAG GCCGAAGCCG TCGCGGCGTA TCTGCAAAAT
 151  ATCGGCGTGA AGTTCGGCGA CACGGTCGCG CTGGCGGTTT CCAATTCCAC
 201  AGAATTTATT ACCGCCTATT TCGCCATCTC CGCCATCGGC GCGGTCGCCG
 251  TACCGATGAA CACATTTTTG AAAAACAGCG AATACGCGTA TATCCTGAAC
 301  GACTGCAAGG CGCGCTTCCT GTTCGCCTCG GCCGGCCTGT CAAAAGAATT
 351  GGCGGGCTTG AAGGCGCAAA CGCCCGTCGA AAAAATCATT TGGACGGACA
 401  AAAGCCGTCC GACCGGCGAA ACGGCGGAAG GCGATGCCTT TTTTGAAGAC
 451  GTGCGCCGCT TCCCCGAAAA ACCCGACTTG GGCCGCCAAC CCCGGATAAA
 501  TGATTTGGCA CACATCATCT ACACCTCCGG CACGACGGGG CATCCCAAAG
 551  GCGCGCTAAT CAGTTACGCC AACCTGTTCG CCAACCTGAA CGGCATCGAA
 601  CGCATCTTTA AAATTTCCAA GCGCGACCGC TTTATCGTTT TCCTGCCGAT
 651  GTTCCACAGC TTCACGCTGA CGGCTATGGT GCTGCTGCCG ATTTATATGG
 701  CGTGTTCGAT TATTTTGGTC AAATCCGTTT TTCCGTTTTC CAACGTTTTG
 751  AAACAGACAC TGCTCAAACG CGCGACCGTG TTTTTGGGCG TACCCGCGAT
 801  TTACACCGCG ATGAGCAAGG CGAAAATCCC TTGGTATTTC AGATGGTTCA
 851  ACCGCATTCG CCTGTTTATC AGCGGCGGCG CGCCTTTGGC GGAACAAACC
 901  ATCCTCGATT TCAAAGCCAA GTTCCCCCGC GCCAAATTGC TGGAAGGCTA
```

```
m164.pep    VEVPRGEVGELIVRGGSVMRGYLNMPAATDETIVNGWLKTGDFVTIDEDGFIFIVDRKKD
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g164        VEVPRGEVGELIVRGGSVMRGYLNMPAATDETIVNGWLKTGDFVTIDEDGFIFIVDRKKD
                    280       290       300       310       320       330

              420       430       440       450       460       470
m164.pep    LIISKGQNVYPREIEEEIYKLDAVEAAAVIGVKDRYADEEIVAFVQLKEGMDLGENEIRR
            ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||:||||
g164        LIISKGQNVYPREIEEEIHKLDAVEAAAVIGVKDRYADEEIVAFVQLKEGMDLGEDEIRR
                    340       350       360       370       380       390

              480       490       500       510
m164.pep    HLRTVLANFKIPKQIHFKDGLPRNATGKVLKRVLKEQFDGNKX
            |||||||||||||||||||||||||||||||||||||||:||||
g164        HLRTVLANFKIPKQIHFKDGLPRNATGKVLKRVLKEQFEGNKX
                    400       410       420       430
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 683>:

```
a164.seq
      1  ATGAACCGGA CTTATGCCAA TTTCTACGAA ATGCTGACCG CCGCCTGCCG
     51  CAAAAACGGC AACGGCACGG CAGTGTTCGA CGGCAAGGAA AAAACCGCCT
    101  ACCGCGCGCT CAAGCAGGAA GCCGAAGCCG TTGCGGCGTA TCTGCAAAAT
    151  ATCGGCGTGA AGTTCGGCGA CACGGTCGCG CTGGCGGTTT CCAATTCCAC
    201  GGAATTTATT ACCGCCTATT TCGCCGTATC CGCCATCGGC GCGGTTGCCG
    251  TACCGATGAA CACATTTTTG AAAAACAGCG AATACGCGTA TATCCTGAAC
    301  GACTGCAAGG CGCGCTTCCT GTTCGCCTCG GCCGGCCTGT CAAAAGAATT
    351  GGCGGGCTTG AAGGCGCAAA CGCCCGTCGA AAAAATCATT TGGACGGGCC
    401  AAAGCCGTCC GGACGGCGAA ATGGCGGAAG GCGATGCCTT TTTTGAAGAC
    451  GTGCGCCGCT TCCCCGAAAA ACCCGACTTG GGCCGCCAAC CCCGGATAAA
    501  TGATTTGGCA CACATCATCT ACACCTCCGG CACGACGGGG CATCCCAAAG
    551  GTGCGCTAAT CAGCTACGCC AACCTGTTCG CCAACCTGAA CGGCATCGAA
    601  CGCATCTTTA AAATCTCCAA GCGCGACCGC TTTATCGTTT TCCTGCCGAT
    651  GTTCCACAGC TTCACGCTGA CGGCTACGGT GCTGCTGCCG ATTTATATGG
    701  CGTGTTCGAT TATTTTGGTC AAATCCGTTT TCCCCTTTTC CAACGTTTTG
    751  AAACAGGCAC TGCTCAAACG CGCGACCGTG TTTTTGGGCG TGCCCGCGAT
    801  TTACACCGCG ATGAGCAAGA CGAAAATCCC TTGGTATTTC AGATGGTTCA
    851  ACCGCATCCG CCTGTTTATC AGCGGCGGAG CACCTTTGGC GGAACAAACC
    901  ATCCTCGATT TCAAAGCCAA GTTCCCCCGC GCCAAATTGC TGGAAGGCTA
    951  CGGACTGAGC GAAGCCTCGC CCGTCGTCGC CGTCAATACG CCCGAGAGGC
   1001  AAAAAGCCCG CAGCGTCGGC ATCCCCCTGC CCGGTTTGGA AGTCAAAGCC
   1051  GTCGATGAAG AATTGGTCGA AGTGCCGCGC GGCGAAGTGG GCGAACTGAT
   1101  CGTCAGGGGC GGTTCGGTGA TGCGGGGCTA CCTCAATATG CCTGCCGCCA
   1151  CCGATGAAAC CATCGTCAAC GGCTGGTTGA AAACGGGCGA TTTCGTTACC
   1201  ATAGACGAAG ACGGCTTTAT CTTTATCGTC GACCGCAAAA AAGATTTGAT
   1251  TATTTCCAAA GGTCAAAATG TCTATCCGCG CGAAATCGAA GAAGAAATCT
   1301  ACAAACTCGA TGCCGTCGAA GCCGCCGCCG TCATCGGCGT GAAAGACCGT
   1351  TATGCCGACG AGGAAATCGT CGCCTTCGTC CAATTGAAGG AAGGTATGGA
   1401  TTTGGGCGAG AACGAAATCC GCCGCCACCT GCGTACCGTG CTGGCAAATT
   1451  TCAAAATCCC CAAACAAATC CACTTTAAAG ACGGGCTGCC GCGCAACGCT
   1501  ACGGGCAAGG TATTGAAACG GGTGTTGAAG GAGCAGTTTG ACGGAAACAA
   1551  ATGA
```

This corresponds to the amino acid sequence <SEQ ID 684; ORF 164.a>:

```
a164.pep
      1  MNRTYANFYE MLTAACRKNG NGTAVFDGKE KTAYRALKQE AEAVAAYLQN
     51  IGVKFGDTVA LAVSNSTEFI TAYFAVSAIG AVAVPMNTFL KNSEYAYILN
    101  DCKARFLFAS AGLSKELAGL KAQTPVEKII WTGQSRPDGE MAEGDAFFED
    151  VRRFPEKPDL GRQPRINDLA HIIYTSGTTG HPKGALISYA NLFANLNGIE
    201  RIFKISKRDR FIVFLPMFHS FTLTAMVLLP IYMACSIILV KSVFPFSNVL
    251  KQALLKRATV FLGVPAIYTA MSKTKIPWYF RWFNRIRLFI SGGAPLAEQT
    301  ILDFKAKFPR AKLLEGYGLS EASPVVAVNT PERQKARSVG IPLPGLEVKA
    351  VDEELVEVPR GEVGELIVRG GSVMRGYLNM PAATDETIVN GWLKTGDFVT
    401  IDEDGFIFIV DRKKDLIISK GQNVYPREIE EEIYKLDAVE AAAVIGVKDR
    451  YADEEIVAFV QLKEGMDLGE NEIRRHLRTV LANFKIPKQI HFKDGLPRNA
```

```
  51  GACTTTGGGC GTTTTGCTCA AAGAACTCGA ACCGTCTTGG GAAATCACCC
 101  TGATTGAACG CTTGGAagat gTGGCGTTGG AATCGTCAAA cGCGTGGAAC
 151  AACGcCGgca CGGGGCATTC CGcGCTGTGc GAATTGAACT AtgcgccGCT
 201  GGGtgcggac ggcgtcatCA ATCCGGCGCg cgCCCTGAAT ATTGCCGAAC
 251  AGTTTCATGT CAGCCGCCAG TTTTGGGcga cgctgGTCGC GGAAGGCAAG
 301  TTGGAagaCA ATTCCTTCAT CAATGCcgtg ccgcatatGT Ctttggtgat
 351  gAacgaagac cactgCCgtt acCTGCAAAA ACGCTATGAT GTGTTTAAAA
 401  CGCAGAAACT TTTTGAAAAT ATGGAATTTT CCACCGATCG GAACAAAATT
 451  TCCGATTGGG CtccgCTGAT TATGCGCGGC CGGgacgaaA ACCAACCCGT
 501  CGCCGCCAAC TATTCCGCCG Aaggcacgga tgtcgATTTC GGACGGCTGA
 551  CGCGCCAGAT GGTGAAATAT TTGCAGGGCA AGGGCGTAAA AACCGAGTTC
 601  AACCGCCACG TCGAAGACAT CAAACGCGAA TCCGACGGCG CGTGGGTGCT
 651  CAAAACCGCC GATACCCGCA ACCCAGACTG GCAGCTCACC CTCCGCACCC
 701  GCTTCCTCTT CCTCGGCGCG GGCGGCGGCG CACTGACCCT GCTGCAAAAA
 751  TCCGGCATCC CCGAAGGCAA AGGCTACGGC GGCTTACCCG TGTCCGGCCT
 801  GTTCTTCCGC AACAGCAACC CCGAAACCGC CGAACAACAC AACGCCAAAG
 851  TGTACGGGCA GGCTTCCGTC GGCGCGCCGC CGATGTCCGT CCCGCACCTC
 901  GACACACGCA ACGTAGACGG CAAACGACAC CTTATGTTCG GTCCTTACGC
 951  AGGTTTCCGT TCCAACTTCC TCAAGCAAGG CTCGTTTATG GATTTGCCGC
1001  TGTCCATCCA TATGGACAAC CTCTATCCTA TGCTGCGCGC CGGCTGGGCG
1051  AATATGCCGC TGACCAAATA CcTGCTGGgC gAaTTGCgtt aa
```

This corresponds to the amino acid sequence <SEQ ID 686; ORF 165.ng>:

```
g165.pep
   1  MAEATDVVLV GGGIMSATLG VLLKELEPSW EITLIERLED VALESSNAWN
  51  NAGTGHSALC ELNYAPLGAD GVINPARALN IAEQFHVSRQ FWATLVAEGK
 101  LEDNSFINAV PHMSLVMNED HCRYLQKRYD VFKTQKLFEN MEFSTDRNKI
 151  SDWAPLIMRG RDENQPVAAN YSAEGTDVDF GRLTRQMVKY LQGKGVKTEF
 201  NRHVEDIKRE SDGAWVLKTA DTRNPDWQLT LRTRFLFLGA GGGALTLLQK
 251  SGIPEGKGYG GLPVSGLFFR NSNPETAEQH NAKVYGQASV GAPPMSVPHL
 301  DTRNVDGKRH LMFGPYAGFR SNFLKQGSFM DLPLSIHMDN LYPMLRAGWA
 351  NMPLTKYLLG ELR*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 687>:

```
m165.seq  (partial)
   1  ATGGCTGAAG CGACAGACGT TGTCTTGGTG GGCGGCGGCA TTATGAGCGC
  51  GACTTTGGGC GTTTTGCTCA AAGAACTCGA ACCGTCTTGG GAAATCACCC
 101  TGATTGAACG CTTGGAAGAT GTGGCGTTGG AATCGTCAAA CGCGTGGAAC
 151  AACGCCGGCA CGGGGCATTC CGCGCTGTGC GAATTGAACT ATGCGCCGTT
 201  GGGTGCAAAT GGGATTATCG ATCCGGCGCG CGCCCTCAAT ATTGCCGAAC
 251  AGTTTCATGT CAGCCGCCAG TTTTGGGCGA CGCTGGTCGC GGAAGGCAAG
 301  TTGGAAGACA ATTCCTTCAT CAATGCCGTG CCGCATATGT CTTTGGTGAT
 351  GAATGAAGAC CATTGTTCTT ATCTTCAAAA ACGTTATGAC GCGTTTAAAA
 401  CCCAAAAACT TTTTGAAAAT ATGGAATTTT CCACCGATCG GAACAAAATT
 451  TCCGATTGGG CTCCGCTGAT GATGCGCGGC CGGGACGAAA ACCAACCCGT
 501  CGCCGCCAAC TACTCCGCCG AAGgTACGGA TGTCGATTTC GGACGGCTGA
 551  CGCGCCAAAT GGTGAAATAT TTGCAGGGCA AGGGCGTAAA AACCGAGTTC
 601  AACCGCCACG TCGAAGACAT CAAACGCGAA TCCGACGGCG CGTGGGTGCT
 651  CAAAACCGCC GATACCCGCA ACCCCGACGG GCAGCTCACC CTCCGTACCC
 701  GCTTCCTCTT CCTCGGCGCG GGCGGCGGCG CGCTGACCCT GCTGCAAAAA
 751  TCCGGCATCC CCGAAGGCAA AGGCTACGGC GGCTTCCCCG TGTCCGGCCT
 801  GTTCTTCCGC AACAGCAACC CCGAAACCGC CGAACAACAC AACGCCAAAG
 851  TGTACGGGCA GGCTTCCGTC GGCGCGCCGC CGATGTCCGT CCCGCACCTC
 901  GACACACGCA ACGTGGACGG CAAACGCCAC CTTATGTTCG GCCCTTACGC
 951  AGGCTTCCGT TCCAACTTCC TCAAGCAAGG CTCGCTTATG GATTTGCCGC
1001  TGTCCATCCA TATGGACAAC CTCTATCCTA TGCTGTGCGC CGGCTGGGCG
1051  AATATGCCGC TGACCAAA...
```

This corresponds to the amino acid sequence <SEQ ID 688; ORF 165>:

```
m165.pep  (partial)
   1  MAEATDVVLV GGGIMSATLG VLLKELEPSW EITLIERLED VALESSNAWN
  51  NAGTGHSALC ELNYAPLGAN GIIDPARALN IAEQFHVSRQ FWATLVAEGK
 101  LEDNSFINAV PHMSLVMNED HCSYLQKRYD AFKTQKLFEN MEFSTDRNKI
 151  SDWAPLMMRG RDENQPVAAN YSAEGTDVDF GRLTRQMVKY LQGKGVKTEF
```

```
 601  AACCGCCACG TCGAAGACAT CAAACGCGAA TCCGACGGCG CGTGGGTGCT
 651  CAAAACCGCC GATACCCGCA ACCCCGACGG GCAGCTCACC CTCCGTACCC
 701  GCTTCCTCTT CCTCGGCGCG GGCGGCGGCG CGCTGACCCT GCTGCAAAAA
 751  TCCGGCATCC CCGAAGGCAA AGGCTACGGT GGCTTTCCCG TGTCCGGCCT
 801  GTTCTTCCGC AACAGCAACC CCGAAACCGC CGAACAACAC AACGCCAAAG
 851  TGTACGGGCA GGCTTCCGTC GGCGCGCCGC CGATGTCCGT CCCGCACCTC
 901  GACACACGCA ACGTGGACGG CAAACGCCAC CTTATGTTCG GCCCTTACGC
 951  AGGCTTCCGT TCCAACTTCC TCAAGCAAGG CTCACTTATG GATTTGCCGC
1001  TGTCCATCCA TATGGACAAC CTCTATCCTA TGCTGCGCGC CGGCTGGGCG
1051  AATATGCCGC TGACCAAATA CCTGCTGGGC GAATTGCGTA AAACCAAAGA
1101  AGAACGCTTC GCCTCCCTGC TGGAATACTA CCCCGAGGCA AACCCCGACG
1151  ACTGGGAACT CATCACCGCA GGGCAACGCG TTCAAATCAT TAAAAAAGAC
1201  TCCGAAAAAG GCGGCGTGTT GCAGTTTGGT ACGGAGATTG TCGCACACGC
1251  CGACGGCTCG CTCGCCGCAT TGCTGGGCGC GTCGCCGGGC GCATCGACCG
1301  CCGTGCCGCT GATGATCCGG CTGATGCACC AATGCTTCCC CGAACGCACC
1351  CCGTCTTGGG AAGGCCGTCT GAAAGAGCTG GTACCGGGTT ACGGCATCAA
1401  GTTGAACGAA AACCCCGAAA GGGCGGATGA AATTATCGCC TATACCGCGA
1451  AAGTGTTGGA TATTTAA
```

This corresponds to the amino acid sequence <SEQ ID 690; ORF 165.a>:

```
a165.pep
   1  MAEATDVVLV GGGIMSATLG VLLKELEPSW EITLIERLED VALESSNAWN
  51  NAGTGHSALC ELNYAPLGAN GIIDPARALN IAEQFHVSRQ FWATLVAEGK
 101  LEDNSFINAV PHMSLVMNED HCSYLQKRYD AFKTQKLFEN MEFSTDRNKI
 151  SDWAPLMMRG RDENQPVAAN YSAEGTDVDF GRLTRQMVKY LQGKGVKTEF
 201  NRHVEDIKRE SDGAWVLKTA DTRNPDGQLT LRTRFLFLGA GGGALTLLQK
 251  SGIPEGKGYG GFPVSGLFFR NSNPETAEQH NAKVYGQASV GAPPMSVPHL
 301  LTRNVDGKRH LMFGPYAGFR SNFLKQGSLM DLPLSIHMDN LYPMLRAGWA
 351  NMPLTKYLLG ELRKTKEERF ASLLEYYPEA NPDDWELITA GQRVQIIKKD
 401  SEKGGVLQFG TEIVAHADGS LAALLGASPG ASTAVPLMIR LMHQCFPERT
 451  PSWEGRLKEL VPGYGIKLNE NPERADEIIA YTAKVLDI*
```

m165/a165   99.7% identity in 356 aa overlap

```
                    10        20        30        40        50        60
m165.pep    MAEATDVVLVGGGIMSATLGVLLKELEPSWEITLIERLEDVALESSNAWNNAGTGHSALC
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a165        MAEATDVVLVGGGIMSATLGVLLKELEPSWEITLIERLEDVALESSNAWNNAGTGHSALC
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m165.pep    ELNYAPLGANGIIDPARALNIAEQFHVSRQFWATLVAEGKLEDNSFINAVPHMSLVMNED
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a165        ELNYAPLGANGIIDPARALNIAEQFHVSRQFWATLVAEGKLEDNSFINAVPHMSLVMNED
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m165.pep    HCSYLQKRYDAFKTQKLFENMEFSTDRNKISDWAPLMMRGRDENQPVAANYSAEGTDVDF
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a165        HCSYLQKRYDAFKTQKLFENMEFSTDRNKISDWAPLMMRGRDENQPVAANYSAEGTDVDF
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m165.pep    GRLTRQMVKYLQGKGVKTEFNRHVEDIKRESDGAWVLKTADTRNPDGQLTLRTRFLFLGA
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a165        GRLTRQMVKYLQGKGVKTEFNRHVEDIKRESDGAWVLKTADTRNPDGQLTLRTRFLFLGA
                   190       200       210       220       230       240

                   250       260       270       280       290       300
m165.pep    GGGALTLLQKSGIPEGKGYGGFPVSGLFFRNSNPETAEQHNAKVYGQASVGAPPMSVPHL
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a165        GGGALTLLQKSGIPEGKGYGGFPVSGLFFRNSNPETAEQHNAKVYGQASVGAPPMSVPHL
                   250       260       270       280       290       300
```

```
 851  TGTACGGGCA GGCTTCCGTC GGCGCGCCGC CGATGTCCGT CCCGCACCTC
 901  GACACACGCA ACGTGGACGG CAAACGCCAC CTTATGTTCG GCCCTTACGC
 951  AGGCTTCCGT TCCAACTTCC TCAAGCAAGG CTCGCTTATG GATTTGCCGC
1001  TGTCCATCCA TATGGACAAC CTCTATCCTA TGCTGTGCGC CGGCTGGGCG
1051  AATATGCCGC TGACCAAATA CCTGCTGGGC GAATTGCGTA AAACCAAAGA
1101  AGAACGCTTC GCCTCCCTGC TGGAATACTA CCCCGAGGCA AACCCCGACG
1151  ACTGGGAACT CATCACCGCA GGGCAACGCG TCCAAATCAT TAAAAAAGAC
1201  TCCGAAAAAG GCGGCGTGCT CCAGTTTGGT ACGGAGATTG TCGCCCACGC
1251  CGACGGCTCA CTCGCCGCAT TGCTGGGCGC GTCGCCGGGC GCATCGACCG
1301  CTGTGCCGCT GATGATCCGG CTGATGCACC AATGCTTCCC CGAGCGCGCC
1351  CCGTCTTGGG AAGACCGTCT GAAAGAGCTG GTACCGGGTT ACGGCATCAA
1401  GTTGAACGAA AACCCTGAAA GGGCGGATGA AATTATCGCC TATACCGCGA
1451  AAGTATTGGA TATTTAA
```

This corresponds to the amino acid sequence <SEQ ID 694; ORF 165-1>:

m165-1.pep

```
  1  MAEATDVVLV GGGIMSATLG VLLKELEPSW EITLIERLED VALESSNAWN
 51  NAGTGHSALC ELNYAPLGAN GIIDPARALN IAEQFHVSRQ FWATLVAEGK
101  LEDNSFINAV PHMSLVMNED HCSYLQKRYD AFKTQKLFEN MEFSTDRNKI
151  SDWAPLMMRG RDENQPVAAN YSAEGTDVDF GRLTRQMVKY LQGKGVKTEF
201  NRHVEDIKRE SDGAWVLKTA DTRNPDGQLT LRTRFLFLGA GGGALTLLQK
251  SGIPEGKGYG GFPVSGLFFR NSNPETAEQH NAKVYGQASV GAPPMSVPHL
301  DTRNVDGKRH LMFGPYAGFR SNFLKQGSLM DLPLSIHMDN LYPMLCAGWA
351  NMPLTKYLLG ELRKTKEERF ASLLEYYPEA NPDDWELITA GQRVQIIKKD
401  SEKGGVLQFG TEIVAHADGS LAALLGASPG ASTAVPLMIR LMHQCFPERA
451  PSWEDRLKEL VPGYGIKLNE NPERADEIIA YTAKVLDI*
```

m165-1/g165-1   89.7% identity in 428 aa overlap

```
                        10        20        30        40        50        60
m165-1.pep      MAEATDVVLVGGGIMSATLGVLLKELEPSWEITLIERLEDVALESSNAWNNAGTGHSALC
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g165-1          MAEATDVVLVGGGIMSATLGVLLKELEPSWEITLIERLEDVALESSNAWNNAGTGHSALC
                        10        20        30        40        50        60

                        70        80        90       100       110       120
m165-1.pep      ELNYAPLGANGIIDPARALNIAEQFHVSRQFWATLVAEGKLEDNSFINAVPHMSLVMNED
                |||||||||:|:|:||||||||||||||||||||||||||||||||||||||||||||||
g165-1          ELNYAPLGADGVINPARALNIAEQFHVSRQFWATLVAEGKLEDNSFINAVPHMSLVMNED
                        70        80        90       100       110       120

                       130       140       150       160       170       180
m165-1.pep      HCSYLQKRYDAFKTQKLFENMEFSTDRNKISDWAPLMMRGRDENQPVAANYSAEGTDVDF
                || |||||||:||||||||||||||||||||||||||:||||||||||||||||||||||
g165-1          HCRYLQKRYDVFKTQKLFENMEFSTDRNKISDWAPLIMRGRDENQPVAANYSAEGTDVDF
                       130       140       150       160       170       180

                       190       200       210       220       230       240
m165-1.pep      GRLTRQMVKYLQGKGVKTEFNRHVEDIKRESDGAWVLKTADTRNPDGQLTLRTRFLFLGA
                |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
g165-1          GRLTRQMVKYLQGKGVKTEFNRHVEDIKRESDGAWVLKTADTRNPDWQLTLRTRFLFLGA
                       190       200       210       220       230       240

                       250       260       270       280       290       300
m165-1.pep      GGGALTLLQKSGIPEGKGYGGFPVSGLFFRNSNPETAEQHNAKVYGQASVGAPPMSVPHL
                |||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
g165-1          GGGALTLLQKSGIPEGKGYGGLPVSGLFFRNSNPETAEQHNAKVYGQASVGAPPMSVPHL
                       250       260       270       280       290       300

                       310       320       330       340       350       360
m165-1.pep      DTRNVDGKRHLMFGPYAGFRSNFLKQGSLMDLPLSIHMDNLYPMLCAGWANMPLTKYLLG
                ||||||||||||||||||||||||||||:||||||||||||||||| ||||||||||||||
g165-1          DTRNVDGKRHLMFGPYAGFRSNFLKQGSFMDLPLSIHMDNLYPMLRAGWANMPLTKYLLG
                       310       320       330       340       350       360

                       370       380       390       400       410       420
m165-1.pep      ELRKTKEERFASLLEYYPEANPDDWELITAGQRVQIIKKDSEKGGVLQFGTEIVAHADGS
                ||||||||||||||||||: :     |||   | :||  || |  |||:   :   |  :
g165-1          ELRKTKEERFASLLEYYPR-QTRRLVLITQXTR-HIIXYDS-KLRVLQLYEIVPRDARSR
                       370        380       390        400       410

                       430       440       450       460       470       480
```

```
a165-1.pep    GRLTRQMVKYLQGKGVKTEFNRHVEDIKRESDGAWVLKTADTRNPDGQLTLRTRFLFLGA
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m165-1        GRLTRQMVKYLQGKGVKTEFNRHVEDIKRESDGAWVLKTADTRNPDGQLTLRTRFLFLGA
                      190       200       210       220       230       240

                      250       260       270       280       290       300
a165-1.pep    GGGALTLLQKSGIPEGKGYGGFPVSGLFFRNSNPETAEQHNAKVYGQASVGAPPMSVPHL
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m165-1        GGGALTLLQKSGIPEGKGYGGFPVSGLFFRNSNPETAEQHNAKVYGQASVGAPPMSVPHL
                      250       260       270       280       290       300

                      310       320       330       340       350       360
a165-1.pep    DTRNVDGKRHLMFGPYAGFRSNFLKQGSLMDLPLSIHMDNLYPMLRAGWANMPLTKYLLG
              |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
m165-1        DTRNVDGKRHLMFGPYAGFRSNFLKQGSLMDLPLSIHMDNLYPMLCAGWANMPLTKYLLG
                      310       320       330       340       350       360

                      370       380       390       400       410       420
a165-1.pep    ELRKTKEERFASLLEYYPEANPDDWELITAGQRVQIIKKDSEKGGVLQFGTEIVAHADGS
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m165-1        ELRKTKEERFASLLEYYPEANPDDWELITAGQRVQIIKKDSEKGGVLQFGTEIVAHADGS
                      370       380       390       400       410       420

                      430       440       450       460       470       480
a165-1.pep    LAALLGASPGASTAVPLMIRLMHQCFPERTPSWEGRLKELVPGYGIKLNENPERADEIIA
              |||||||||||||||||||||||||||||:|||| |||||||||||||||||||||||||
m165-1        LAALLGASPGASTAVPLMIRLMHQCFPERAPSWEDRLKELVPGYGIKLNENPERADEIIA
                      430       440       450       460       470       480

                     489
a165-1.pep    YTAKVLDIX
              |||||||||
m165-1        YTAKVLDIX
```

a165-1/p33940

```
sp|P33940|YOJH_ECOLI HYPOTHETICAL 60.2 KD PROTEIN IN ECO-ALKB INTERGENIC REGION
>gi|1736851|gnl|PID|d1016718 (D90850) ORF_ID:o372#5; similar to [SwissProt Accession Number
P33940] [Escherichia coli] >gi|1788539 (AE000310) f548; This 548 aa ORF is 100 pct identical
to 490 residues of YOJH_ECOLI SW: P33940 (492 aa) but contains 56 additional N-ter aa; 100 pct
identical to GB: ECOHU49_33
ACCESSION: U00008 (490 aa) but contains 58 aditional N-term resi... Length = 548
 Score =  458 bits (1167), Expect = e-128
 Identities = 233/490 (47%), Positives = 303/490 (61%), Gaps = 5/490 (1%)

Query: 3   EATDVVLVGGGIMSATLGVLLKELEPSWEITLIERLEDVALESSNAWNNAGTGHSALCEL 62
            + TDV+L+GGGIMSATLG  L+ELEP W +T++ERLE VA ESSN WNNAGTGHSAL EL
Sbjct: 30  QETDVLLIGGGIMSATLGTYLRELEPEWSMTMVERLEGVAQESSNGWNNAGTGHSALMEL 89

Query: 63  NYAPLGANGIIDPARALNIAEQFHVSRQFWATLVAEGKLED-NSFINAVPHMSLVMNEDH 121
            NY P  A+G I    +A+ I E F +SRQFWA  V  G L    SFIN VPHMS V  ED+
Sbjct: 90  NYTPQNADGSISIEKAVAINEAFQISRQFWAHQVERGVLRTPRSFINTVPHMSFVWGEDN 149

Query: 122 CSYLQKRYDAFKTQKLFENMEFSTDRNKISDWAPLMMRGRDENQPVAANYSAEGTDVDFG 181
            ++L+ RY A +   LF  M +S D  +I +WAPL+M GRD  Q VAA  +  GTDV++G
Sbjct: 150 VNFLRARYAALQQSSLFRGMRYSEDHAQIKEWAPLVMEGRDPQQKVAATRTEIGTDVNYG 209

Query: 182 RLTRQMVKYLQGKG-VKTEFNRHVEDIKRESDGAWVLKTADTRNPDGQLTXXXXXXXXXX 240
            +TRQ++  LQ K      + +  V  +KR  D W +  AD +N   Q
Sbjct: 210 EITRQLIASLQKKSNFSLQLSSEVRALKRNDDNTWTVTVADLKNGTAQ-NIRAKFVFIGA 268

Query: 241 XXXXXXXXQKSGIPEGKGYGGFPVSGLFFRNSNPETAEQHNAKVYGQASVGAPPMSVPHL 300
                    Q+SGIPE K Y GFPV G F  + NP+    H AKVYG+ASVGAPPMSVPH+
Sbjct: 269 GGAALKLLQESGIPEAKDYAGFPVGGQFLVSENPDVVNHHLAKVYGKASVGAPPMSVPHI 328

Query: 301 DTRNVDGKRHLMFGPYAGFRSNFLKQGSLMDLPLSIHMDNLYPMLRAGWANMPLTKYLLG 360
            DTR +DGKR ++FGP+A F + FLK GSL DL  S    N+ PM+  G  N  L KYL+
Sbjct: 329 DTRVLDGKRVVLFGPFATFSTKFLKNGSLWDLMSSTTTSNVMPMMHVGLDNFDLVKYLVS 388

Query: 361 ELRKTKEERFASLLEYYPEANPDDWELITAGQRVQIIKKDSEKGGVLQFGTEIVXXXXXX 420
            ++   ++E+RF +L EYYP+A   +DW L  AGQRVQIIK+D+EKGGVL+ GTE+V
Sbjct: 389 QVMLSEEDRFEALKEYYPQAKKEDWRLWQAGQRVQIIKRDAEKGGVLRLGTEVVSDQQGT 448

Query: 421 XXXXXXXXXXXXXXVPLMIRLMHQCFPER--TPSWEGRLKELVPGYGIKLNENPERADEI 478
```

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 204 shows 82.0% identity over a 250 aa overlap with a predicted ORF (ORF 204.ng) from *N. gonorrhoeae:*

```
m204/g204

                       10        20        30        40        50        60
m204.pep       MAAAEIKRPFAVDFQHIASVLHGGIAAFACLIGLQGGMRNXVIRQFAAVFGDIAHQFGKQ
               |||||||||:|||||||||||||||||||||||||||||| || ||||||||||||||||
g204           MAAAEIKRPLAVDFQHIASVLHGGIAAFACLIGLQGGMRNQVISQFAAVFGDIAHQFGKQ
                       10        20        30        40        50        60

                       70        80        90       100       110       120
m204.pep       GMAHAVFCPTCRTVLIIGFHTFAADGFQVAGMLADQSDNILFRQAFNRITDLFFAVVGFA
               ||||||| |: | || :|||||| |||||:|||: | |::||||||||||||||||||||
g204           GMAHAVFRPARRRVLSVGFHTFADDGFQVVGMLSGQPDGVLFRQAFNRITDLFFAVVGFA
                       70        80        90       100       110       120

                      130       140       150       160       170       180
m204.pep       FAALSQIQTGNRRIVDIYDFENRFRRALYRVLRLYRRIXGFAATAXQQAAAQYGKXXXQH
               ||:||| |||||||||::|||||||||| |:|||:||| |||| : ||||||:||   ||
g204           FATLSQSQTGNRRIVDVFDFENRFRRALCRILRLFRRIFGFAAGGKQQAAAQHGKRYFQH
                      130       140       150       160       170       180

                      190           200       210       220       230
m204.pep       STXLMVSKCRLK----RGRRRFGRHRVHFNGRMPTASGTLSNNSRASLRAFAAPACKISS
               |: |||||||||    ||||||||| |:||||||||| ||||||||||||| ||||||||
g204           SALLMVSKCRLKCRLKRGRRRFGRHWVYFNGRMPTASRTLSNNSRASLRAFCAPACKISS
                      190       200       210       220       230       240

               240
m204.pep       ICEGSAVSSLX
               ||||  |::|
g204           ICEGLEVNAL
                      250
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 701>:

```
a204.seq
      1  ATGGCGGCGG CGGAAATAAA ACGCCCCCTC GCTGTCGATT TCCAGCACAT
     51  AGCGTCCGTT CTGCACGGCG GCATAGCCGC TTTTGCCTGC CTGATAGGGT
    101  TGCAGGGCGG AATGCGAAAT CAGGTAATCC GTCAGTTTGC CGCCGTCTTC
    151  GGCGATATTG CCCACCAGTT TGGCAAACAA GGTATGGCAC ACGCCGTTTG
    201  CCGCCCAGCC CGAAGGCGCG CCCTTTCCGT CGGTTTCCAT ACATTTGCCG
    251  ACGACGGCTT CCAAGTCGTT GGGATGCTTG CCGGTCAGCC GGACGACGTT
    301  TTGTTCCGGC AAGCCTTT.. .......... .......... ..........
    351  .......... .......... .......... .......... ..........
    401  .......... .......... .......... .......... ..........
    451  .......... .......... .......... .......... ..........
    501  .......... .......... .......... .......... ..........
    551  .......... .......... .......... .....AAGAG GTTCGGACGG
    601  CATTGGGTTT ATTTCAACGG GCGGATACCG ACCGCATCAC GTACTTTGCC
    651  CAATAATTCG CGTGCTTCTT TACGCGCTTT TTGCGCGCCT GCCTGCAAAA
    701  TCTCTTCGAT TTGCGAAGGG TCGGCGGTCA GCTCGTTGTA G
```

This corresponds to the amino acid sequence <SEQ ID 702; ORF 204.a>:

```
a204.pep
      1  MAAAEIKRPL AVDFQHIASV LHGGIAAFAC LIGLQGGMRN QVIRQFAAVF
     51  GDIAHQFGKQ GMAHAVCRPA RRRALSVGFH TFADDGFQVV GMLAGQPDDV
    101  LFRQAF.... .......... .......... .......... ..........
    151  .......... .......... .......... .......... .....KRFGR
    201  HWVYFNGRIP TASRTLPNNS RASLRAFCAP ACKISSICEG SAVSSL*
```

This corresponds to the amino acid sequence <SEQ ID 706; ORF 205>:

```
m205.pep
     1  MLXTXFAVLG GCLLXCRCGK SXNTAVQPQN AVQSAPKPVF KVIYIDNTAI
    51  AGLDLGQSSE GKTNDGKKQI SYPIKGLPEQ NVIRLIGKHP GDLEAVSGKC
   101  METDDKDSPA GWAENGVCHT LFAKLVGNIA EDGGKLTDYL VSHAALQPYQ
   151  AGKSGYAAVQ NGRYVLEIDS EGAFYFRRRH Y*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 205 shows 88.4% identity over a 181 aa overlap with a predicted ORF (ORF 205.ng) from *N. gonorrhoeae:*

```
m205/g205

                    10        20        30        40        50        60
m205.pep   MLXTXFAVLGGCLLXCRCGKSXNTAVQPQNAVQSAPKPVFKVIYIDNTAIAGLDLGQSSE
           ||  |||||||||  |||| ||| |||||:|||||||||| ||||||||| ||||||
g205       MLKIPFAVLGGCLLLAACGKSENTAEQPQNAAQSAPKPVFKVKYIDNTAIAGLALGQSSE
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m205.pep   GKTNDGKKQISYPIKGLPEQNVIRLIGKHPGDLEAVSGKCMETDDKDSPAGWAENGVCHT
           ||||||||||||||||||||::|| ||||:||||| |||||| ||:|:|||||||||
g205       GKTNDGKKQISYPIKGLPEQNAVRLTGKHPNDLEAVVGKCMETDGKDAPSGWAENGVCHT
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m205.pep   LFAKLVGNIAEDGGKLTDYLVSHAALQPYQAGKSGYAAVQNGRYVLEIDSEGAFYFRRRH
           ||||||||||||||||||||:||:||||||||||||||||||||||||||||||||||||
g205       LFAKLVGNIAEDGGKLTDYLISHSALQPYQAGKSGYAAVQNGRYVLEIDSEGAFYFRRRH
                   130       140       150       160       170       180

m205.pep   YX
           |
g205       Y
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 707>:

```
a205.seq  (partial)
     1  TCCGAACCTC TTAAAGGCTT GCCGGAACAA AACGTCGTCC GGCTGACCGG
    51  CAAGCATCCC AACGACTTGG AAGCCGTCGT CGGCAAATGT ATGGAAACCG
   101  ACGGAAAGGG CGCGCCTTCG GGCTGGGCGG CAAACGGCGT GTGCCATACC
   151  TTGTTTGCCA AACTGGTGGG CAATATCGCC GAAGACGGCG GCAAACTGAC
   201  GGATTACCTG ATTTCGCATT CCGCCCTGCA ACCCTATCAG GCAGGCAAAA
   251  GCGGCTATGC CGCCGTGCAG AACGGACGCT ATGTGCTGGA AATCGACAGC
   301  GAGGGGGCGT TTTATTTCCG CCGCCGCCAT TATTGA
```

This corresponds to the amino acid sequence <SEQ ID 708; ORF 205.a>:

```
a205.pep  (partial)
     1  SEPLKGLPEQ NVVRLTGKHP NDLEAVVGKC METDGKGAPS GWAANGVCHT
    51  LFAKLVGNIA EDGGKLTDYL ISHSALQPYQ AGKSGYAAVQ NGRYVLEIDS
   101  EGAFYFRRRH Y*
```

m205/a205 88.3% identity in 111 aa overlap

```
                    50        60        70        80        90       100
m205.pep   KVIYIDNTAIAGLDLGQSSEGKTNDGKKQISYPIKGLPEQNVIRLIGKHPGDLEAVSGKC
                                       | |:||||||||:|| ||||:||||| |||
a205                                   SEPLKGLPEQNVVRLTGKHPNDLEAVVGKC
                                                10        20        30

                   110       120       130       140       150       160
m205.pep   METDDKDSPAGWAENGVCHTLFAKLVGNIAEDGGKLTDYLVSHAALQPYQAGKSGYAAVQ
```

```
                      70        80        90       100       110       120

                     130       140       150       160       170
g205-1.pep     LFAKLVGNIAEDGGKLTDYLISHSALQPYQAGKSGYAAVQNGRYVLEIDSEGAF
               ||||||||||||||||||||:||:||||||||||||||||||||||||||||||
m205-1         LFAKLVGNIAEDGGKLTDYLVSHAALQPYQAGKSGYAAVQNGRYVLEIDSEGAFYFRRRH
                     130       140       150       160       170       180

m205-1         YX
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 713>:

a205-1.seq (partial)

```
  1  CCTCTTAAAG GCTTGCCGGA ACAAAACGTC GTCCGGCTGA CCGGCAAGCA
 51  TCCCAACGAC TTGGAAGCCG TCGTCGGCAA ATGTATGGAA ACCGACGGAA
101  AGGGCGCGCC TTCGGGCTGG GCGGCAAACG GCGTGTGCCA TACCTTGTTT
151  GCCAAACTGG TGGGCAATAT CGCCGAAGAC GGCGGCAAAC TGACGGATTA
201  CCTGATTTCG CATTCCGCCC TGCAACCCTA TCAGGCAGGC AAAAGCGGCT
251  ATGCCGCCGT GCAGAACGGA CGCTATGTGC TGGAAATCGA CAGCGAGGGG
301  GCGTTTTATT TCCGCCGCCG CCATTATTGA
```

This corresponds to the amino acid sequence <SEQ ID 714; ORF 205-1.a>:

a205-1.pep (partial)

```
  1  PLKGLPEQNV VRLTGKHPND LEAVVGKCME TDGKGAPSGW AANGVCHTLF
 51  AKLVGNIAED GGKLTDYLIS HSALQPYQAG KSGYAAVQNG RYVLEIDSEG
101  AFYFRRRHY*
```

m205-1/a205-1   89.0% identity in 109 aa overlap

```
                    50        60        70        80        90       100
m205-1.pep     KYIDNTAIAGLDLGQSSEGKTNDGKKQISYPIKGLPEQNVIRLIGKHPGDLEAVSGKCME
                                             |:||||||||:|| ||||:||||| |||||
a205-1                                       PLKGLPEQNVVRLTGKHPNDLEAVVGKCME
                                                     10        20        30

                   110       120       130       140       150       160
m205-1.pep     TDDKDSPAGWAENGVCHTLFAKLVGNIAEDGGKLTDYLVSHAALQPYQAGKSGYAAVQNG
               || | :|:||| |||||||||||||||||||||||||||:||:|||||||||||||||||
a205-1         TDGKGAPSGWAANGVCHTLFAKLVGNIAEDGGKLTDYLISHSALQPYQAGKSGYAAVQNG
                     40        50        60        70        80        90

                   170       180
m205-1.pep     RYVLEIDSEGAFYFRRRHYX
               ||||||||||||||||||||
a205-1         RYVLEIDSEGAFYFRRRHYX
                     100       110
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 715>:

g206.seq

```
  1  atgttttccc ccgacaaaac ccttttcctc tgtctcggcg cactgctcct
 51  cgcctcatgc ggcacgacct ccggcaaaca ccgccaaccg aaacccaaac
101  agacagtccg gcaaatccaa gccgtccgca tcagccacat cggccgcaca
151  caaggctcgc aggaactcat gctccacagc ctcggactca tcggcacgcc
201  ctacaaatgg ggcggcagca gcaccgcaac cggcttcgac tgcagcggca
251  tgattcaatt ggtttacaaa aacgccctca acgtcaagct gccgcgcacc
301  gcccgcgaca tggcggcggc aagccgcaaa atccccgaca gccgcctcaa
351  ggccggcgac atcgtattct tcaacaccgg cggcgcacac cgctactcac
401  acgtcggact ctacatcggc aacggcgaat tcatccatgc ccccggcagc
451  ggcaaaacca tcaaaaccga aaaactctcc acaccgtttt acgccaaaaa
501  ctaccttgga gcgcatacgt tttttacaga atga
```

This corresponds to the amino acid sequence <SEQ ID 716; ORF 206.ng>:

g206.pep

```
  1  MFSPDKTLFL CLGALLLASC GTTSGKHRQP KPKQTVRQIQ AVRISHIGRT
```

This corresponds to the amino acid sequence <SEQ ID 720; ORF 206.a>:

```
a206.pep
      1  MFPPDKTLFL CLSALLLASC GTTSGKHRQP KPKQTVRQIQ AVRISHIDRT
     51  QGSQELMLHS LGLIGTPYKW GGSSTATGFD CSGMIQFVYK NALNVKLPRT
    101  ARDMAAASRK IPDSRLKAGD LVFFNTGGAH RYSHVGLYIG NGEFIHAPSS
    151  GKTIKTEKLS TPFYAKNYLG AHTFFTE*
```

m206/a206 99.4% identity in 177 aa overlap

```
                        10        20        30        40        50        60
m206.pep        MFPPDKTLFLCLSALLLASCGTTSGKHRQPKPKQTVRQIQAVRISHIDRTQGSQELMLHS
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a206            MFPPDKTLFLCLSALLLASCGTTSGKHRQPKPKQTVRQIQAVRISHIDRTQGSQELMLHS
                        10        20        30        40        50        60

                        70        80        90       100       110       120
m206.pep        LGLIGTPYKWGGSSTATGFDCSGMIQFVYKNALNVKLPRTARDMAAASRKIPDSRXKAGD
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
a206            LGLIGTPYKWGGSSTATGFDCSGMIQFVYKNALNVKLPRTARDMAAASRKIPDSRLKAGD
                        70        80        90       100       110       120

                       130       140       150       160       170
m206.pep        LVFFNTGGAHRYSHVGLYIGNGEFIHAPSSGKTIKTEKLSTPFYAKNYLGAHTFFTEX
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a206            LVFFNTGGAHRYSHVGLYIGNGEFIHAPSSGKTIKTEKLSTPFYAKNYLGAHTFFTEX
                       130       140       150       160       170
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 721>:

```
g209.seq
      1  atgctgcggc atttaggaaa cgacttcgcc ttgggcgcgt tgtttttcga
     51  tgctgcggtt gatgtgccac tgctgggcga tggtcaggag gttgttgacc
    101  acccagtaga gaaccaaacc ggcagggaag aagaagaaca tgacggagaa
    151  aaccaacggc atgattttca tcattttcgc ctgcatcggg tcggtcggcg
    201  gcgggttcag ataggtttgg gcgaacatcg ttgccgccat aatgatgggc
    251  aggatgtagt aggggtcggc gcggctgagg tcggtaatcc agcccagcca
    301  aggtgcctgg cgcaattcta cggaggcgaa caatgcccag tacaagccga
    351  tgaagacggg gatttgcaac agcataggca gacagccgcc cagcgggttg
    401  atttcctcgt cttcgaaaag ctgcatcatc gcttgctgtt gcgccatacg
    451  gtcgtcgccg tatttttctt tgatggtctg cagttcgggt gcggcggcac
    501  gcatttcgc catcgaacgg taggaggcgt tggtcaatgg atacagtacg
    551  gctttgacga tgatggtcaa aacgacgatt gcccagcccc agttgccgat
    601  aatgttgtgc agttggttca ggagccagaa gagcggcgat gcgaaccagt
    651  gtactttacc gtagtctttt gccagttgca ggttgtcggc gatgtttgcg
    701  ataacggatg tggtttgcgg accggcatac aggttgaccg ccatttttcgg
    751  ttttggcccc cgggttggga tagcggttaa
```

This corresponds to the amino acid sequence <SEQ ID 722; ORF 209.ng>:

```
g209.pep
      1  MLRHLGNDFA LGALFFDAAV DVPLLGDGQE VVDHPVENQT GREEEEHDGE
     51  NQRHDFHHFR LHRVGRRRVQ IGLGEHRCRH NDGQDVVGVG AAEVGNPAQP
    101  RCLAQFYGGE QCPVQADEDG DLQQHRQTAA QRVDFLVFEK LHHRLLLRHT
    151  VVAVFFFDGL QFGCGGTHFR HRTVGGVGQW IQYGFDDDGQ NDDCPAPVAD
    201  NVVQLVQEPE ERRCEPVYFT VVFCQLQVVG DVCDNGCGLR TGIQVDRHFR
    251  FWPPGWDSG*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 723>:

```
m209.seq
      1  ATGCTGCGGC ATTTAGGAAA CGACTTCGCC TTGGgGCGTT GTTTTTCGAT
     51  GCTGCGGTTG ATGTGCCATT GCTGGGCGAT GGTCAGGAGG TTGTTGACTA
    101  CCCAGTACAA TACCAGACCG GCAGGGAAGA AGAAGAACAT GACGGAGAAA
    151  ACCAACGGCA TGATTTTCAT CATTTTCGCC TGCATCGGGT CGGTCGGCGG
    201  CGGGTTCAGA TAAGTTTGGG CGAACATCGT TGCCGCCATA ATGATGGGCA
    251  GGATGTAGTA GGGGTCGGCG CGGCTGAGGT CGGTAATCCA ACCCAGCCAA
```

```
  51  TGCTGCGGTT GATGTGCCAT TGCTGGGCGA TGGTCAGGAG GTTGTTGATC
 101  ACCCAGTACA ATACCAGACC GGCAGGGAAG AAGAAGAACA TGACGGAGAA
 151  AACCAAAGGC ATGATTTTCA TCATTTTCGC CTGCATCGGG TCGGTCGGCG
 201  GCGGGTTCAG ATAGGTTTGG GCGAACATCG TTGCCGCCAT AATGATGGGC
 251  AGGATGTAGT AGGGGTCGGC GCGGCTGAGG TCGGTAATCC AACCCAGCCA
 301  AGGTGCCTGG CGCAATTCTA CGGAGGCGAA CAATGCCCAA TACAATCCGA
 351  TGAAGACGGG GATTTGCAAC AGCATAGGCA GGCAGCCGCC CAGCGGGTTG
 401  ATTTTCTCGT CTGTGTAAAG CTGCATCATG GCTTGTTGCT GCGCCATACG
 451  GTCGTCGCCG TATTTCTCTT TGATGGCTTG CAGTTTGGGC GCGGCGGCAC
 501  GCATTTTCGC CATCGAACGG TAAGAGGCGT TGGTCAATGG ATACAGTACG
 551  GCTTTGACGA TGATGGTTAA AACGATAATC GCCCAGCCCC AGTTGCCGAT
 601  GATGTTGTGC AGTTGGTTCA AAAGCCAAAA GAGGGGGGAG GCGAACCAGT
 651  GTACTTTGCC GTAGTCTTTG GCCAGTTGCA GGTTGTCGGC GATGTTTGCG
 701  ATAACGGATG TGGTCTGTGG GCCGGCGTAG AGGTTGATGG AGGCTTCGGT
 751  TTCGCACCGT TTTGGATAGC GGCTAAAGGC ACGCTGACGC TGGTGCTGTA
 801  CAGCTTGTCG TTGCGGCGTT TGATGTCGAT ACGGCAGTCG CCAGCGGCGC
 851  AAACGCTTTG TCCGCCTTTT GGTTGGAGGA TCCAGGTGGA CATGAAGTGG
 901  TGTTCAATCA TGCCGAGCCA GCCGGTCGGG GTTTTGCGGA TGTATTCGGC
 951  CTCGGATTTG CCGGATTTGG CATCGTCGTC CAAGTCGGAG AAGCTGACTT
1001  TTTGGAAGTT GCCTTCAGGG GTATAA
```

This corresponds to the amino acid sequence <SEQ ID 726; ORF 209.a>:

```
a209.pep
   1  MLRHLGNDFA LGALFFDAAV DVPLLGDGQE VVDHPVQYQT GREEEEHDGE
  51  NQRHDFHHFR LHRVGRRRVQ IGLGEHRCRH NDGQDVVGVG AAEVGNPTQP
 101  RCLAQFYGGE QCPIQSDEDG DLQQHRQAAA QRVDFLVCVK LHHGLLLRHT
 151  VVAVFLFDGL QFGRGGTHFR HRTVRGVGQW IQYGFDDDG* NDNRPAPVAD
 201  DVVQLVQKPK EGGGEPVYFA VVFGQLQVVG DVCDNGCGLW AGVEVDGGFG
 251  FAPFWIAAKG TLTLVLYSLS LRRLMSIRQS PAAQTLCPPL GWRIQVDMKW
 301  CSIMPSQPVG VLRMYSASDL PDLASSSKSE KLTFWKLPSG V*
```

m209/a209 95.6% identity in 341 aa overlap

```
                    10        20        30        40        50        60
m209.pep    MLRHLGNDFALGALFFDAAVDVPLLGDGQEVVDYPVQYQTGREEEEHDGENQRHDFHHFR
            ||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
a209        MLRHLGNDFALGALFFDAAVDVPLLGDGQEVVDHPVQYQTGREEEEHDGENQRHDFHHFR
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m209.pep    LHRVGRRRVQISLGEHRCRHNDGQDVVGVGAAEVGNPTQPRCLAQFYGGEQCPIQSDEDG
            ||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
a209        LHRVGRRRVQIGLGEHRCRHNDGQDVVGVGAAEVGNPTQPRCLAQFYGGEQCPIQSDEDG
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m209.pep    DLQQHRQAAAQRVDFLVCVKLHHRLLLRHTVVAVFLFDGLQFGCGGTHFRHRAVRGVGQW
            ||||||||||||||||||||||| |||||||||||||||||| ||||||||:||||||||
a209        DLQQHRQAAAQRVDFLVCVKLHHGLLLRHTVVAVFLFDGLQFGRGGTHFRHRTVRGVGQW
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m209.pep    IQYGFDDDGXNDNRPAPVADDVVQLVQEPEERGGEPVYFAVVFGQLQVVGDVCDDGCGLR
            |||||||||||||||||||||||||||:|:| ||||||||||||||||||||||:|||| 
a209        IQYGFDDDGXNDNRPAPVADDVVQLVQKPKEGGGEPVYFAVVFGQLQVVGDVCDNGCGLW
                   190       200       210       220       230       240

                   250       260       270       280       290       300
m209.pep    AGVEVDGGFGFAPFWMAAKGTLTLVLYSLSLRRLMSMLHSPAAQTLCLPLGWRIQVDMKW
            |||||||||||||||:||||||||||||||||||||: :|||||||| ||||||||||||
a209        AGVEVDGGFGFAPFWIAAKGTLTLVLYSLSLRRLMSIRQSPAAQTLCPPLGWRIQVDMKW
                   250       260       270       280       290       300

                   310       320       330       340
m209.pep    CSIMPSQPVGVLRMYSASDLPDLASSSKSEKLTFWKLPSGVX
```

```
               ||||:||||||||||||||||||||||||||||||||::|||:|||||||||::||| :|
g211           AFVVLQREVTFFGEDDVVAAVFVLFAQEQVFHDAGFGVKGFDEINPAVALAQVIELAGFH
                     70        80        90       100       110       120

                    130       140       150       160       170
m211.pep       QRQFLLLLQDFSVFAAAXLCPRYHPKLHDGNQNGKRHGKLHHRAYPLFQCQSAGX
               |||||||||:|::|||| ||||||||||||||:||||||||  |||||| ||||
g211           QRQFLLLLQNFGIFAAAQLCPRYHPKLHDGNQDGKRHGKLHDGAYPLFQRQSAG
                    130       140       150       160       170
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 731>:

```
a211.seq
      1  ATGTTGCGGG TTGCTGCTGC CAATCAGTTG GGCGGTCGGA ATGGTACGGC
     51  GGTGGGAAAC GGGGTCGATG AGTTTGGGCG TGGTGCTGAT AATCAGGTTG
    101  AGTTTTTGGA AGGAAACCTG ATTGTAGTCG GCGCGTCCGG GCGTGCCGCT
    151  GTAACGGTAG CCGTGGCGCA ATTCGAGCGT GCGTTTGTTG TCGTTCAGCG
    201  AGAAGTTACT TTCTTTGGCG AAGATGATGT TGTCGCCGCC GTTTTTGTCC
    251  TGTTCGCGCA GGAACAGGTT TTTCATGATG CCGGATTCGG TATCGAAGGT
    301  TTCGACAAAA TAAACCCTGC CGTTGCGCTT GCCCAAACTG TTGAACCCGC
    351  CTGCCTCCAC CAAAGACAAT TCCTGCTTCT GCTTCAGGAT TTCAGCGTAT
    401  TCGCGGCTGC GTAGCTCTGC CCACGGTATC ACCCAAAGCT GCATGACGGC
    451  AACCAAAACG GCAAACGGCA CGGCAAACTG CATCACCGGG CGTATCCATT
    501  GTTTCAATGC CAATCCGCAG GATAG
```

This corresponds to the amino acid sequence <SEQ ID 732; ORF 211.a>:

```
a211.pep
      1  MLRVAAANQL GGRNGTAVGN GVDEFGRGAD NQVEFLEGNL IVVGASGRAA
     51  VTVAVAQFER AFVVVQREVT FFGEDDVVAA VFVLFAQEQV FHDAGFGIEG
    101  FDKINPAVAL AQTVEPACLH QRQFLLLLQD FSVFAAA*LC PRYHPKLHDG
    151  NQNGKRHGKL HHRAYPLFQC QSAG*
```

m211/a211 99.4% identity in 174 aa overlap

```
                     10        20        30        40        50        60
m211.pep       MLRVAAANQLGGRNGTAVGNGVDEFGRGADNQVEFLEGNLIVVGASGRAAVTVAVAQFER
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a211           MLRVAAANQLGGRNGTAVGNGVDEFGRGADNQVEFLEGNLIVVGASGRAAVTVAVAQFER
                     10        20        30        40        50        60

                     70        80        90       100       110       120
m211.pep       AFVVVQREVTFFGEDDVVAAVFVLFAQEQVFHDAGFGIEGFDKINPAVALAQTVELACLH
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
a211           AFVVVQREVTFFGEDDVVAAVFVLFAQEQVFHDAGFGIEGFDKINPAVALAQTVEPACLH
                     70        80        90       100       110       120

                    130       140       150       160       170
m211.pep       QRQFLLLLQDFSVFAAAXLCPRYHPKLHDGNQNGKRHGKLHHRAYPLFQCQSAGX
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
a211           QRQFLLLLQDFSVFAAAXLCPRYHPKLHDGNQNGKRHGKLHHRAYPLFQCQSAGX
                    130       140       150       160       170
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 733>:

```
g212.seq (partial)
      1  atggacaatc tcgtatggga cggcattccc gacatccgca cactcgacca
     51  aaccatccgc aaacacgcac acccgctcaa cctgattgtc tgcctccccg
    101  ataatcagat tcccgatttt caaaccgcac aagatgcttc ggactcggaa
    151  tgccgtctga agcaccgttt ggatcaggca acccagtgcc tccagttcga
    201  cagcatcaac ctcatcgaac acatcctgcc cgatgtccgc ttctggctgg
    251  ttccccctttc acgcacccgc cgcctgcacg aacacttcca ccacatttcc
    301  tggcagaccg aagccatccc gcaaaccgaa agcaagtccg acaaaccctg
    351  gtttgcactt ccacaaacat ccgaacggaa aaaaccggaa cacgtcctcg
```

```
1501  GmCAwTGCAG CCCAAATCsT AGGCyTGCCC CATCCCTTTT yAcAAcGCCT
1551  gCGCCACGCC cTAcACCCCA ACCGCACCAT CATCCGCGCC ATCGTCAGAA
1601  GGAAGGATCT AACCCCTTAA
```

This corresponds to the amino acid sequence <SEQ ID 736; ORF 212>:

```
m212.pep
     1  MDNLVWDGIP DIRTLDQAIR KHAPPLNLII CLPDNQIPDF QTAQDASDAE
    51  CRLKHRLDQA MQCLQFDSIN LIEHILPDVR FWLVPPSRTH HLHEHFHHIS
   101  WQTEAIPQTE SKPDKPWFAL PQTSERQKPE HILVIGAGIS GAATAHALAS
   151  HGISVTVLEA RKAAQAASGN RQGLLYAKIS PHDTEQTELL LAGYGYTKRL
   201  LGHILPESET WGGNGIIHLN YSRTEQQRNH ELGLQKHHNH LYRSITSAEA
   251  EKIAGIPLSV PYDHPSCGLY WQHGVWLNPP AFVRTLLNHP LIGLHEDTPL
   301  TDISHDGEKW IASTPNGTFT ATHIIYCTGA NSPYLPETNL AALPLRQIRG
   351  QTGLTPSTPF SEQLRCAVSG ESYISPSWHG LHCYGASFIP NSSHTGWNEA
   401  EEASNRQALA HLNPALSESL FAANPNPQKH QGHAAIRCDS PDHLPLVGAL
   451  GDIAAMRQTY TKLALDKNYR IDTPCPYLPN AYVNTAHGTR GLATAPICAA
   501  XXAAQIXGLP HPFXQRLRHA LHPNRTIIRA IVRRKDLTP*
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

ORF 212 shows 92.9% identity over a 421 aa overlap with a predicted ORF (ORF 212.ng) from *N. gonorrhoeae:*

```
m212/g212

                    10        20        30        40        50        60
m212.pep    MDNLVWDGIPDIRTLDQAIRKHAPPLNLIICLPDNQIPDFQTAQDASDAECRLKHRLDQA
            ||||||||||||||||:||||| |||||:|||||||||||||||||||:||||||||||
g212        MDNLVWDGIPDIRTLDQTIRKHAHPLNLIVCLPDNQIPDFQTAQDASDSECRLKHRLDQA
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m212.pep    MQCLQFDSINLIEHILPDVRFWLVPPSRTHHLHEHFHHISWQTEAIPQTESKPDKPWFAL
             ||||||||||||||||||||||||||||::|||||||||||||||||||||| ||||||
g212        TQCLQFDSINLIEHILPDVRFWLVPPSRTRRLHEHFHHISWQTEAIPQTESKSDKPWFAL
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m212.pep    PQTSERQKPEHILVIGAGISGAATAHALASHGISVTVLEARKAAQAASGNRQGLLYAKIS
            ||||||:||||:|||||||:||:|||||||||||||||||||||||||||||||||||||
g212        PQTSERKKPEHVLVIGAGIAGASTAHALASHGISVTVLEARKAAQAASGNRQGLLYAKIS
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m212.pep    PHDTEQTELLLAGYGYTKRLLGHILPESETWGGNGIIHLNYSRTEQQRNHELGLQKHHNH
            |||| |||||||||||||||||||||:|:|||||||||||||||||||||||||||||||
g212        PHDTGQTELLLAGYGYTKRLLGHILPDSDTWGGNGIIHLNYSRTEQQRNHELGLQKHHNH
                   190       200       210       220       230       240

                   250       260       270       280       290       300
m212.pep    LYRSITSAEAEKIAGIPLSVPYDHPSCGLYWQHGVWLNPPAFVRTLLNHPLIGLHEDTPL
            ||||||||||||||||||::|| :| ||||||||||||||||||||||:||||  |:|:| |
g212        LYRSITSAEAEKIAGIPLNTPYAEPLCGLYWQHGVWLNPPAFVRTLLSHPLIELYENTTL
                   250       260       270       280       290       300

                   310       320       330       340       350       360
m212.pep    TDISHDGEKWIASTPNGTFTATHIIYCTGANSPYLPETNLAALPLRQIRGQTGLTPSTPF
            | ||||||||||||||||||||||||||||:|| |||||||||||||||||||||||||
g212        TGISHDGEKWIASTPNGTFTATHIIYCTGAHSPCLPETNLAALPLRQIRGQTGLTPSTPF
                   310       320       330       340       350       360

                   370       380       390       400       410       420
m212.pep    SEQLRCAVSGESYISPSWHGLHCYGASFIPNSSHTGWNEAEEASNRQALAHLNPALSESL
```

```
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m212.pep    MQCLQFDSINLIEHILPDVRFWLVPPSRTHHLHEHFHHISWQTEAIPQTESKPDKPWFAL
            |||||||||||||||||||||||||||||::|||||||||||||||||||||||||||||
a212        TQCLQFDSINLIEHILPDVRFWLVPPSRTRRLHEHFHHISWQTEAIPQTESKPDKPWFAL
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m212.pep    PQTSERQKPEHILVIGAGISGAATAHALASHGISVTVLEARKAAQAASGNRQGLLYAKIS
            ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
a212        PQTSERQKPEHILVIGAGISGAATAHALASYGISVTVLEARKAAQAASGNRQGLLYAKIS
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m212.pep    PHDTEQTELLLAGYGYTKRLLGHILPESETWGGNGIIHLNYSRTEQQRNHELGLQKHHNH
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a212        PHDTEQTELLLAGYGYTKRLLGHILPESETWGGNGIIHLNYSRTEQQRNHELGLQKHHNH
                   190       200       210       220       230       240

                   250       260       270       280       290       300
m212.pep    LYRSITSAEAEKIAGIPLSVPYDHPSCGLYWQHGVWLNPPAFVRTLLNHPLIGLHEDTPL
            ||||||:|||||||||||  ::||  :| |||:||:|||||||:|||:||:|||||||||||
a212        LYRSITQAEAEKIAGIPLNTPYAEPLCGLFWQYGVWLNPPTFVRALLSHPLIGLHEDTPL
                   250       260       270       280       290       300

                   310       320       330       340       350       360
m212.pep    TDISHDGEKWIASTPNGTFTATHIIYCTGANSPYLPETNLAALPLRQIRGQTGLTPSTPF
            ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
a212        TDISHDGEKWIASTPNGTFTATHIIYCTGANSPYLPETNLATLPLRQIRGQTGLTPSTPF
                   310       320       330       340       350       360

                   370       380       390       400       410       420
m212.pep    SEQLRCAVSGESYISPSWHGLHCYGASFIPNSSHTGWNEAEEASNRQALAHLNPALSESL
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a212        SEQLRCAVSGESYISPSWHGLHCYGASFIPNSSHTGWNEAEEASNRQALAHLNPALSESL
                   370       380       390       400       410       420

                   430       440       450       460       470       480
m212.pep    FAANPNPQKHQGHAAIRCDSPDHLPLVGALGDIAAMRQTYTKLALDKNYRIDTPCPYLPN
            ||||||||||||||||||||||||||||||||||||:|||:|||||||||||||:|||||||
a212        FAANPNPQKHQGHAAIRCDSPDHLPLVGALGDIAAMQQTYAKLALDKNYRIDAPCPYLPN
                   430       440       450       460       470       480

                   490       500       510       520       530       540
m212.pep    AYVNTAHGTRGLATAPICAAXXAAQIXGLPHPFXQRLRHALHPNRTIIRAIVRRKDLTPX
            ||:||||||||||||||||||  ||:| |||||: :|||||||||:||||||||||||||
a212        AYANTAHGTRGLATAPICAAAVAAEILGLPHPLSKRLRHALHPNRAIIRAIVRRKDLTPX
                   490       500       510       520       530       540
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 739>:

```
g214.seq
      1  atgatacaaa agatatgtaa gctatttgtt ttaattgtaa tttttgcaac
     51  ttctcccgct tttgcccttc aaagcgacag cagacggccc atccaaatcg
    101  aagccgacca aggttcgctc gatcaagcca accaaaggac cacatttagc
    151  ggcaatgtca tcatcagaca gggtacgctc aacatttccg cctcgtgtgt
    201  caacgtcaca cgcggcaggc aaaggcggcg aatccgtgag ggcggaaggt
    251  tcgcccgtcc gcttcagcca aacgttggac gggggcaaag ggacggtgcg
    301  cggtcaggca aacaacgtta cctattcctc cgcaggaagc actgtcgttc
    351  tgaccggcaa tgccaaagtg cagcgcggcg gcgacgttgc cgaaggtgcg
    401  gtcattacct acaacaccaa aaccgaagtc tataccatca acggcagcac
    451  gaaatcgggt gcgaaatccg cttccaaaac cggcagggtc agcgtcgtca
    501  tccagccttc aagcacacaa aaaaccgaat aaccccgatg ccgtctgaaa
    551  cggaaacgca gttcagacgg catttgccga ccgaaatgcc gagaagagat
    601  tattga
```

This corresponds to the amino acid sequence <SEQ ID 740; ORF 214.ng>:

```
g214.pep
```

```
351  TAACCGGTAA TGCCAAAGTA CAGCGCGGCG GCGATGTCGC CGAAGGTGCG
401  GTGATTACAT ACAACACCAA AACCGAAGTC TATACCATCA GCGGCAGCAC
451  AAAATCCGGC GCAAAATCCG CTTCCAAATC CGGCAGGGTC AGCGTCGTTA
501  TCCAGCCTTC GAGTACGCAA AAATCCGAAT AATCCCAATG CCGTCTGAAA
551  CATAAACCTG GTTCGGACGG CATTTGCCGA CCGAAATATT GAAGAGATAT
601  TTATGA
```

This corresponds to the amino acid sequence <SEQ ID 744; ORF 214.a>:

```
a214.pep
  1  MIQKICKLFV LIAFFSASPA FALQSDSRQP IQIEADQGSL DQANQSTTFS
 51  GNVVIRQGTL NISAARVNVT RGXQRRRIRE GGRFASPLQP DIGRRQRHGA
101  RTGKQRCLFI CRQHRSLNR* CQSTARRRCR RRCGDYIQHQ NRSLYHQRQH
151  KIRRKIRFQI RQGQRRYPAF EYAKIRIIPM PSET*TWFGR HLPTEILKRY
201  L*
```

m214/a214 99.3% identity in 152 aa overlap

```
                    10        20        30        40        50        60
m214.pep    MIQKICKLFVLIAFFSASPAFALQSDSRQPIQIEADQGSLDQANQSTTFSGNVVIRQGTL
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a214        MIQKICKLFVLIAFFSASPAFALQSDSRQPIQIEADQGSLDQANQSTTFSGNVVIRQGTL
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m214.pep    NISAARVNVTRGRQRRRIREGGRFASPLQPDIGRRQRHGARTGKQRCLFICRQHRSLNRX
            ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
a214        NISAARVNVTRGXQRRRIREGGRFASPLQPDIGRRQRHGARTGKQRCLFICRQHRSLNRX
                    70        80        90       100       110       120

                   130       140       150
m214.pep    CQSTARRRCRRRCGDYIQHQNRSLYHQRQHKI
            ||||||||||||||||||||||||||||||||
a214        CQSTARRRCRRRCGDYIQHQNRSLYHQRQHKIRRKIRFQIRQGQRRYPAFEYAKIRIIPM
                   130       140       150       160       170       180

a214        PSETXTWFGRHLPTEILKRYLX
                   190       200
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 745>:

g214-1.seq

```
  1  ATGATACAAA AGATATGTAA GCTATTTGTT TTAATTGTAA TTTTTGCAAC
 51  TTCTCCCGCT TTTGCCCTTC AAAGCGACAG CAGACGGCCC ATCCAAATCG
101  AAGCCGACCA AGGTTCGCTC GATCAAGCCA ACCAAAGTAC CACATTTAGC
151  GGCAATGTCA TCATCAGACA GGGTACGCTC AACATTTCCG CCTCGCGCGT
201  CAACGTCACA CGCGGCGGCA AAGGCGGCGA ATCCGTGAGG GCGGAAGGTT
251  CGCCCGTCCG CTTCAGCCAA ACGTTGGACG GGGGCAAAGG GACGGTGCGC
301  GGTCAGGCAA ACAACGTTAC CTATTCCTCC GCAGGAAGCA CCGTCGTTCT
351  GACCGGCAAT GCCAAAGTGC AGCGCGGCGG CGACGTTGCC GAAGGTGCGG
401  TCATTACCTA CAACACCAAA ACCGAAGTCT ATACCATCAA CGGCAGCACG
451  AAATCGGGTG CGAAATCCGC TTCCAAAACC GGCAGGGTCA GCGTCGTCAT
501  CCAGCCTTCA AGCACACAAA AAACCGAATA A
```

This corresponds to the amino acid sequence <SEQ ID 746; ORF 214-1.ng>:

g214-1.pep

```
  1  MIQKICKLFV LIVIFATSPA FALQSDSRRP IQIEADQGSL DQANQSTTFS
 51  GNVIIRQGTL NISASRVNVT RGGKGGESVR AEGSPVRFSQ TLDGGKGTVR
101  GQANNVTYSS AGSTVVLTGN AKVQRGGDVA EGAVITYNTK TEVYTINGST
151  KSGAKSASKT GRVSVVIQPS STQKTE*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 747>:

m214-1.seq

```
  1  ATGATACAAA AGATATGTAA GCTATTTGTT TTAATAGCAT TTTTTTCGGC
 51  GTCCCCCGCT TTTGCCCTTC AAAGCGACAG CAGGCAGCCT ATTCAGATTG
101  AGGCCGACCA AGGTTCGCTC GATCAAGCCA ACCAAAGCAC CACATTCAGC
151  GGAAACGTCG TCATCAGACA GGGTACGCTC AATATTTCCG CCGCCCGCGT
```

```
 51  GNVVIRQGTL NISAARVNVT RGGKGGESVR AEGSPVRFSQ TLDGGKGTVR
101  GQANNVAYSS AGSTVVLTGN AKVQRGGDVA EGAVITYNTK TEVYTISGST
151  KSGAKSASKS GRVSVVIQPS STQKSE*
```

a214-1/m214-1 100.0% identity in 176 aa overlap

```
                    10        20        30        40        50        60
a214-1.pep  MIQKICKLFVLIAFFSASPAFALQSDSRQPIQIEADQGSLDQANQSTTFSGNVVIRQGTL
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m214-1      MIQKICKLFVLIAFFSASPAFALQSDSRQPIQIEADQGSLDQANQSTTFSGNVVIRQGTL
                    10        20        30        40        50        60

                    70        80        90       100       110       120
a214-1.pep  NISAARVNVTRGGKGGESVRAEGSPVRFSQTLDGGKGTVRGQANNVAYSSAGSTVVLTGN
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m214-1      NISAARVNVTRGGKGGESVRAEGSPVRFSQTLDGGKGTVRGQANNVAYSSAGSTVVLTGN
                    70        80        90       100       110       120

                   130       140       150       160       170
a214-1.pep  AKVQRGGDVAEGAVITYNTKTEVYTISGSTKSGAKSASKSGRVSVVIQPSSTQKSEX
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m214-1      AKVQRGGDVAEGAVITYNTKTEVYTISGSTKSGAKSASKSGRVSVVIQPSSTQKSEX
                   130       140       150       160       170
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 751>:

```
g215.seq
  1  atgaaagtaa gatggcggta cggaattgcg ttcccattga tattggcggt
 51  tgccttgggc agcctgtcgg catggttggg ccgtatcagc gaagtcgaaa
101  tcgaggaagt caggctcaat cccgacgaac ctcaatacac aatggacggc
151  ttggacggaa ggcggtttga cgaacaggga tacttgaaag aacatttgag
201  cgcgaaaggt gcgaaacagt ttcccgaaaa cagcgacatc cattttgatt
251  cgccgcatct cgtgttcttc caagaaggca ggctgttgta cgaagtcggc
301  agcgatgaag ccgtttacca taccgaaaac aaacaggttc tttttaaaaa
351  caacgttgtg ctgaccaaaa ccgccgacgg caggcggcag gcgggtaaag
401  tcgaaaccga aaaactgcac gtcgataccg aatctcaata tgcccaaacc
451  gatacgcctg tcagtttcca atatggcgcg tcgcacggtc aggcgggcgg
501  tatgacctac aaccacaaaa caggcatgtt gaacttctca tctaaagtga
551  aagccgcgat ttatgataca aaagatatgt aa
```

This corresponds to the amino acid sequence <SEQ ID 752; ORF 215.ng>:

```
g215.pep
  1  MKVRWRYGIA FPLILAVALG SLSAWLGRIS EVEIEEVRLN PDEPQYTMDG
 51  LDGRRFDEQG YLKEHLSAKG AKQFPENSDI HFDSPHLVFF QEGRLLYEVG
101  SDEAVYHTEN KQVLFKNNVV LTKTADGRRQ AGKVETEKLH VDTESQYAQT
151  DTPVSFQYGA SHGQAGGMTY NHKTGMLNFS SKVKAAIYDT KDM*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 753>:

```
m215.seq (partial)
  1  ..AGCCTGTCGG CATGGTTGGG TCGTATCAGC GAAGTCGAGA TTGAAGAAGT
 51    CAGGCTCAAT CCCGACGAAC CGCAATACAC AATGGACAGC TTGGACGGCA
101    GGCGGTTTGA CGAACAGGGA TACTTGAAAG AACATTTGAG CGCGAAGGGC
151    GCGAAACAGT TTCCGGAAAG CAGCGACATC CATTTTGATT CGCCGCATCT
201    CGTGTTCTTC CAAGAAGGCA GGTTGTTGTA CGAAGTCGGC AGCGACGAAG
251    CCGTTTACCA TACCGAAAAC AAACAGGTTC TTTTTAAAAA CAACGTTGTG
301    CTGACCAAAA CCGCCGACGG CAAACGGCAG GCGGGTAAAG TTGAAGCCGA
351    AAAGCTGCAC GTCGATACCG AATCTCAATA TGCCCAAACC GATACGCCTG
401    CAGTTTCCA ATATGGTGCA TCGCACGGTC AGGCGGGCGG CATGACTTAC
451    GACCACAwwA CAGGCATGTT GAACTTCTCA TCTAAAGTGA AAGCCACGAT
501    TTATGATACA AAAGATATGT AA
```

This corresponds to the amino acid sequence <SEQ ID 754; ORF 215>:

```
m215.pep  (partial)
  1  ..SLSAWLGRIS EVEIEEVRLN PDEPQYTMDS LDGRRFDEQG YLKEHLSAKG
 51    AKQFPESSDI HFDSPHLVFF QEGRLLYEVG SDEAVYHTEN KQVLFKNNVV
101    LTKTADGKRQ AGKVEAEKLH VDTESQYAQT DTPVSFQYGA SHGQAGGMTY
```

```
               ||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
a215           YLKEHLSSKGAKQFPESSDIHFDSPHLVFFQEGRLLYEVGSDEAVYHTENKQVLFKNNVV
                       70        80        90       100       110       120

                      110       120       130       140       150       160
m215.pep       LTKTADGKRQAGKVEAEKLHVDTESQYAQTDTPVSFQYGASHGQAGGMTYDHXTGMLNFS
               |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
a215           LTKTADGKRQAGKVEAEKLHVDTESQYAQTDTPVSFQYGASHGQAGGMTYDHKTGMLNFS
                      130       140       150       160       170       180

                      170
m215.pep       SKVKATIYDTKDMX
               ||||||||||||||
a215           SKVKATIYDTKDMX
                      190
```

The following partial DNA sequence was identified in *N gonorrhoeae* <SEQ ID 757>:

```
g216.seq (partial)
      1  ..atgatatcga tttcgagctc ggtacccagc gacgaaatca ccgccatcat
     51    ccccgcactc aaacgcaaag acattaccct cgtctgcatc accgcccgcc
    101    ccgattcaac catggcgcgc catgccgata tccacatcac cgcatcggtt
    151    tcgcaagaag cctgcccgtt ggggcttgcc ccgaccacca gcaccaccgc
    201    cgttatggct ttgggcgacg cgttggcggt cgtcctgctg cgcgcccgcg
    251    cgttcacgcc cgacgacttc gccttgatcc accctgccgg cagcctcggc
    301    aaacgcctgc ttttgcgcgt tgccgacatt atgcacaaag gcggcggcct
    351    gcccgccgtc cgactcggca cgcccttgaa aggagccatc gtcagcatga
    401    gcgagaaagg tttgggcatg tgggcgggaa cggacgggca aaggctgtct
    451    gaaaggcctt tttactga
```

This corresponds to the amino acid sequence <SEQ ID 758; ORF 216.ng>:

```
g216.pep (partial)
      1  ..MISISSSVPS DEITAIIPAL KRKDITLVCI TARPDSTMAR HADIHITASV
     51    SQEACPLGLA PTTSTTAVMA LGDALAVVLL RARAFTPDDF ALIHPAGSLG
    101    KRLLLRVADI MHKGGGLPAV RLGTPLKGAI VSMSEKGLGM WAGTDGQRLS
    151    ERPFY*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 759>:

```
m216.seq
      1  ATGGCAATGG CAGAAAACGG AAAATATCTC GACTGGGCAC GCGAAGTGTT
     51  GCACGCCGAA GCGGAAGGCT TGCGCGAAAT TGCAGCGGAA TTGsACAAAA
    101  ACTTCGTCCT TGCGGCAGAC GCGTTGTTGC ACTGCAAGGG CAGGGTCGTT
    151  ATCACGGGCA TGGTCAAGTC GGGACATATC GGGCGCAAAA TGGCGGCAAC
    201  TATGGCCTCG ACCGGCACGC CTGCGTTTTT CGTCCACCCT GCGGAAGCGG
    251  CACACGgCGA TTTGGGTATG ATTGTGGACA rCGACGTGGT CGTCGCGATT
    301  TCCAATTCCG GCGAAAGCGA CGAAATCGCC GCCATCATCC CCGCACTCAA
    351  ACGCAAAGAC ATCACGCTTG TCTGCATCAC CGCCCGCCCC GATTCAACCA
    401  TGGCGCGCCA TGCCGACATC CACATCACGG CGTCGGTTTC CAAAGAAGCC
    451  TGCCCGCTGG GGCTTGCCCC GACCACCAGC ACCACCGCCG TCATGGCTTT
    501  GGGCGATGCG TTGGCGGTCG TCCtGCTGCG CgcACGCGCG TTCACGCCCG
    551  ACGATTTCGC CTTGAGCCAT CCTGCCGGCA GCCTCGGCAA ACGCCTACTT
    601  TTGCGCGTTG CCGACATTAT GCACAAAGGC GGCGGCCTGC CTGCCGTCCG
    651  ACTCGGCACG CCCTTGAAAG AAGCCATCGT CAGCATGAGT GAAAAAGGGC
    701  TGGGCATGTT GGCGGTAACG GACGGGCAAG GCCGTCTGAA AGGCGTATTC
    751  ACCGACGGCG ATTTGCGCCG CCTGTTTCAA GAATGCGACA ATTTTACCGG
    801  TCTTTCGATA GACGAAGTCA TGCATACGCA TCCTAAAACC ATCTCCGCCG
    851  AACGTCTCGC CACCGAAGCC CTGAAAGTCA TGCAGGCAAA CCATGTGAAC
    901  GGGCTTCTGG TTACCGATGC AGATGGCGTG CTGATCGGCG CGCTGAATAT
    951  GCACGACCTG CTGGCGGCAC GGATTGTATA G
```

This corresponds to the amino acid sequence <SEQ ID 760; ORF 216>:

```
m216.pep
      1  MAMAENGKYL DWAREVLHAE AEGLREIAAE LXKNFVLAAD ALLHCKGRVV
```

```
301  GLLVTDADGV LIGALNMHDL LAARIV*
```

m216/a216 97.2% identity in 326 aa overlap

```
                       10        20        30        40        50        60
m216.pep       MAMAENGKYLDWAREVLHAEAEGLREIAAELXKNFVLAADALLHCKGRVVITGMVKSGHI
               |||| | |||||||||||:|||||||||||:| :||:||||||||||||||||||| |||||
a216           MAMAGNEKYLDWAREVLHTEAEGLREIAADLDENFALAADALLHCKGRVVITGMGKSGHI
                       10        20        30        40        50        60

                       70        80        90       100       110       120
m216.pep       GRKMAATMASTGTPAFFVHPAEAAHGDLGMIVDXDVVVAISNSGESDEIAAIIPALKRKD
               |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
a216           GRKMAATMASTGTPAFFVHPAEAAHGDLGMIVDNDVVVAISNSGESDEIAAIIPALKRKD
                       70        80        90       100       110       120

                      130       140       150       160       170       180
m216.pep       ITLVCITARPDSTMARHADIHITASVSKEACPLGLAPTTSTTAVMALGDALAVVLLRARA
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a216           ITLVCITARPDSTMARHADIHITASVSKEACPLGLAPTTSTTAVMALGDALAVVLLRARA
                      130       140       150       160       170       180

                      190       200       210       220       230       240
m216.pep       FTPDDFALSHPAGSLGKRLLLRVADIMHKGGGLPAVRLGTPLKEAIVSMSEKGLGMLAVT
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a216           FTPDDFALSHPAGSLGKRLLLRVADIMHKGGGLPAVRLGTPLKEAIVSMSEKGLGMLAVT
                      190       200       210       220       230       240

                      250       260       270       280       290       300
m216.pep       DGQGRLKGVFTDGDLRRLFQECDNFTGLSIDEVMHTHPKTISAERLATEALKVMQANHVN
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a216           DGQGRLKGVFTDGDLRRLFQECDNFTGLSIDEVMHTHPKTISAERLATEALKVMQANHVN
                      250       260       270       280       290       300

                      310       320
m216.pep       GLLVTDADGVLIGALNMHDLLAARIVX
               |||||||||||||||||||||||||||
a216           GLLVTDADGVLIGALNMHDLLAARIVX
                      310       320
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 763>:

```
g217.seq
      1  atggcggatg acggtttgtt gcggcaactg tccgaaaaac ccagccaaag
     51  tgctctcttc ctgccatttg acccattcgt tttcgaggtt ttggactgcc
    101  ttttggtcat cgggcccggc ttgaaacaat gtttcaagca aatcccggca
    151  acgcgccacc cattcgccga ccgtcgcagg ttgccgccat atccgggcaa
    201  tatccgacag ggtttcgagg aaggcggcaa aacgtccgaa catggcggtt
    251  tgattcacgt cggcatacca cgcgctgaca tcctgccaca tcgggttgcc
    301  gccttcgggc agcatccagc ccaatatcat acggtctgcc gcctgcttcc
    351  aggtaaacag ctgatccgtg ccgccgcgca tttctccgtc caatccccaa
    401  tggacgttca aatcggcaac catatcgtgc aaaagcggca aatcgtcccc
    451  ggtcagtccg aaacggcgca acacgggcgc ggtttccaaa agcgcgagca
    501  ctttgccgac ttcaaaacgg ctttccagca agtcggacac gcactccaac
    551  gcataaaaaa acggttgccg gcggctgatt ttcacgtccg aaacggaata
    601  cggcaatgcc tgcgcgccgg gttgcgcctg tccgaacacg gcttccataa
    651  aaggcgtata gggttcgata ttcggggtta a
```

This corresponds to the amino acid sequence <SEQ ID 764; ORF 217.ng>:

```
g217.pep..
      1  MADDGLLRQL SEKPSQSALF LPFDPFVFEV LDCLLVIGPG LKQCFKQIPA
     51  TRHPFADRRR LPPYPGNIRQ GFEEGGKTSE HGGLIHVGIP RADILPHRVA
    101  AFGQHPAQYH TVCRLLPGKQ LIRAAAHFSV QSPMDVQIGN HIVQKRQIVP
    151  GQSETAQHGR GFQKREHFAD FKTAFQQVGH ALQRIKKRLP AADFHVRNGI
```

```
201  TATCCGTCAG GGTTTCGAGG AAGGCGGCAA AACGTCCGAA CAGGGCGGTT
251  TGGTTCACGT CGGCATACCA CGCGCTGACC CCCTGCCACA TCGGATTGCC
301  GCCTTCGGGC AGCATCCAGC CCAATATCAT GCGTTCTACC GCCTGCTTCC
351  AGGTGAACAG CTGATCCGTG CCGCCGCGCA TTTCTCCGTC CAAACCCCAG
401  CGGACGTTCA AATCGGCAAC CATGTCGTGC AAAAGCGGCA AATCGTCCTC
451  AGTCAGTCCG AAATGGCGCA ACACGGGCGC GGTTTCTAAA AGCACAAGCA
501  CTTTATCGAC TTCAAATCGG CTTTCCAACA AGTCGAACAG GCATGACAAA
551  GCATGAAACA GCGGTTGTCG GCGGCTGATT TTCACATCCG AAACGGAATA
601  CGGCAATGCC TGCGCGCCGG GCTGCGCCTG TCCGAACACG GCTTCGATAA
651  AAGGCGTATA GGATTCGATA TTCGGGGTTA A
```

This corresponds to the amino acid sequence <SEQ ID 768; ORF 217.a>:

a217.pep

```
  1  VADDGVQRQL SGKLRQFGFR LPFDPFVFEA LDCLLVIAFD LEQCFKQIPA
 51  TRHPFVNRRR LPPYPYNIRQ GFEEGGKTSE QGGLVHVGIP RADPLPHRIA
101  AFGQHPAQYH AFYRLLPGEQ LIRAAAHFSV QTPADVQIGN HVVQKRQIVL
151  SQSEMAQHGR GF*KHKHFID FKSAFQQVEQ A*QSMKQRLS AADFHIRNGI
201  RQCLRAGLRL SEHGFDKRRI GFDIRG*
```

m217/a217 90.3% identity in 226 aa overlap

```
                    10        20        30        40        50        60
m217.pep    MADDGVRRQLSGKLRQFGFRLPFDPFVFKVLDXLLVIGFSLEQCFKQIPATRHPFADRCG
            :|||||:|||||||||||||||||||||||::|| ||||:|:||||||||||||||||::|
a217        VADDGVQRQLSGKLRQFGFRLPFDPFVFEALDCLLVIAFDLEQCFKQIPATRHPFVNRRR
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m217.pep    LPPYPYNIRQGFEEGGKTSEHGGLIHVGIPRADILPHRIAAFGQHPAQYHAFYRLLPGEQ
            ||||||||||||||||||||:|||:||||||||| |||||||||||||||||||||||||
a217        LPPYPYNIRQGFEEGGKTSEQGGLVHVGIPRADPLPHRIAAFGQHPAQYHAFYRLLPGEQ
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m217.pep    LIRAAAHFSVQTPVDVQIGNHVVQKRXIVLSQSETAQHGRGFXKHKHFIDFKSAFQQVEQ
            |||||||||||||:|||||||||||| |||||| ||||||||||||||||||||||||||
a217        LIRAAAHFSVQTPADVQIGNHVVQKRQIVLSQSEMAQHGRGFXKHKHFIDFKSAFQQVEQ
                   130       140       150       160       170       180

                   190       200       210       220
m217.pep    AXQSMKQRLAAADFHVXHGIRQCLRTGLRLSEHGFDKRRIGFDIRGX
            |||||||||:|||||: :||||||:||||||||||||||||||||||
a217        AXQSMKQRLSAADFHIRNGIRQCLRAGLRLSEHGFDKRRIGFDIRGX
                   190       200       210       220
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 769>:

g218.seq

```
  1  atggttgcgg tggatcctta tacggcaaaa gtggtcaaca ccatgccgcg
 51  caatcagggt tggtatcaca ctatggatga aatccacggc gatatgatgc
101  tcggtgcggc aggcgattat cttttggaaa cggcagcttc actgaccatt
151  attatggttg tcagcggctt gtacctttgg tgggcgaaac agcgcggcat
201  taaagcgatg ctgctgccgc caaaaagcag ggcgcgttct tggtggcgga
251  atctgcacgg cgcgtttgga acttgggtgt cgttgatttt actgttgttc
301  tgcctgtcgg gtattgcttg ggcaggtatt tggggcggca aattcgtgca
351  ggcttggaat cagttcccgg ccggcaaatg ggtgtcgaa ccgaaccccg
401  tttcaatcgt gccgacccac ggcgaggtat tgaatgacgg caaggttaag
451  gaagtgccgt ggattttgga gcttatgcct atgcctgtct cagggacgac
501  tgtgggtgaa aacggcatta accccaccga gcccaataac attggaaacc
551  gtcgaccgtt tcgcgcggga aatcggtttc aaagggcgtt atcagttgaa
601  tttgcccaaa ggcgaggacg ggtatggac tttgtcgcag gattctatga
651  gttatga
```

This corresponds to the amino acid sequence <SEQ ID 770; ORF 218.ng>:

g218.pep

```
  1  MVAVDPYTAK VVNTMPRNQG WYHTMDEIHG DMMLGAAGDY LLETAASLTI
```

a218.seq

```
  1   ATGGTCGCGG TCGATCCTTA TACGGCAAAA GTGGTCAGTA CCATGCCGCG
 51   CAATCAGGGT TGGTATTACG CGATGGATGA AATCCACAGC GATATGATGC
101   TCGGTTCGAC AGGTGATTAT CTTTTGGAAA CGGCTGCATC GCTGACGATT
151   ATCATGATAA TCAGCGGTTT GTACCTTTGG TGGGTGAAAC GGCGCGGCAT
201   CAAGGCGATG CTGCTGCCGC CAAAAGGCAG GGCGCGTTCT TGGTGGCGGA
251   ATCTGCACGG CGCGTTTGGA ACTTGGGTGT CGTTGATTTT ACTGTTGTTC
301   TGCCTGTCGG GTATTGCTTG GGCAGGTATT TGGGGCGGCA AGTTCGTGCA
351   GGCTTGGAGT CAGTTCCCGG CAGGCAAATG GGGTGTCGAA CCGAACCCTG
401   TTTCAGTCGT GCCGACCCAC GGCGAGGTAT TGAATGACGG CAAGGTTAAG
451   GAAGTGCCGT GGGTTTTGGA GCTTACGCCT ATGCCTGTTT CAGGGACGAC
501   TGTGGGCAAA GACGGTATTA ACCCTGACGA GCCGATGACA TTGGAAACCG
551   TCGACCGTTT TGCGCGG.GA AATCGGTTTC AAAGGGCGTT ATCAGCTGAA
601   TTTGCCCAAA GGCGAGGACG GCGTATGGAC TTTGTCGCAG GATTCTATGA
651   GTTA
```

This corresponds to the amino acid sequence <SEQ ID 774; ORF 218.a>:

a218.pep

```
  1   MVAVDPYTAK VVSTMPRNQG WYYAMDEIHS DMMLGSTGDY LLETAASLTI
 51   IMIISGLYLW WVKRRGIKAM LLPPKGRARS WWRNLHGAFG TWVSLILLLF
101   CLSGIAWAGI WGGKFVQAWS QFPAGKWGVE PNPVSVVPTH GEVLNDGKVK
151   EVPWVLELTP MPVSGTTVGK DGINPDEPMT LETVDRFARX NRFQRALSAE
201   FAQRRGRRMD FVAGFYEL
```

m218/a218 95.9% identity in 218 aa overlap

```
                      10        20        30        40        50        60
m218.pep      MVAVDPYTAKVVSTMPRNQGWYYTMDEIHSDMMLGAAGDYLLETAASLTIIMVVSGLYLW
              |||||||||||||||||||||||:||||||||||::||||||||||||||||::||||||
a218          MVAVDPYTAKVVSTMPRNQGWYYAMDEIHSDMMLGSTGDYLLETAASLTIIMIISGLYLW
                      10        20        30        40        50        60

                      70        80        90       100       110       120
m218.pep      WVKRRGIKAMLLPSKGXARSWWRNLHGTFGTWVSLILLLFCLSGIAWAGIWGGKFVQAWS
              ||||||||||||| || ||||||||||:||||||||||||||||||||||||||||||||
a218          WVKRRGIKAMLLPPKGRARSWWRNLHGAFGTWVSLILLLFCLSGIAWAGIWGGKFVQAWS
                      70        80        90       100       110       120

                     130       140       150       160       170       180
m218.pep      QFPAGKWGVEPNPVSVVPTHGEVLNDGKVKEVPWVLELTPMPVSGTTVGKDGINPDEPMT
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a218          QFPAGKWGVEPNPVSVVPTHGEVLNDGKVKEVPWVLELTPMPVSGTTVGKDGINPDEPMT
                     130       140       150       160       170       180

                     190       200       210
m218.pep      LETVDRFARXNRFQRALSVEFAQRRGRRMDFVAGFYEL
              ||||||||||||||||||:|||||||||||||||||||
a218          LETVDRFARXNRFQRALSAEFAQRRGRRMDFVAGFYEL
                     190       200       210
```

```
                        70        80        90       100       110       120

                       130       140       150       160       170       180
m219.pep        LGWWSVLANVLFCLAVIFIGISGCVMWWKRRPTGAVGIVPPAQKVKLPVWWMMALPLLAI
                ||||||||||:|||||||||||||||||||||:|::||||||||:|||||| |||||| |
g219            LGWWSVLANVVFCLAVIFIGISGCVMWWKRRPSGVAGIVPPAQKIKLPVWWAMALPLLLI
                       130       140       150       160       170       180

                       190       200       210
m219.pep        ALLFPTSLLAIAVIWLLDTLLLSRIPVLRRWFKX
                ||||||:||||||||||||||||||||||:|||
g219            ALLFPTALLAIAVIWLLDTLLLSRIPVLRKWFK
                       190       200       210
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 779>:

```
a219.seq
      1  ATGACGGCAA GGTTAAGGAA GTGCCGTGGG TTTTGGAGCT TACGCCTATG
     51  CCTGTTTCAG GGACGACTGT GGGCAAAGAC GGTATTAACC CTGACGAGCC
    101  GATGACATTG GAAACCGTCG ACCGTTTTGC GCGG.GAAAT CGGTTTCAAA
    151  GGGCGTTATC AGCTGAATTT GCCCAAAGGC GAGGACGGCG TATGGACTTT
    201  GTCGCAGGAT TCTATGAGTT ACGACATCAT CAGCCCGTTT GCTGACCGCA
    251  CGGTGCATAT CGACCAGTAC AGCGGCAAGA TTCTTGCCGA CATCCGTTTT
    301  GACGATTACA ACCCGTTCGG CAAATTTATG GCC[illegible]AGCA TTGCGCTGCA
    351  TATGGGGACT TTGGGCTGGT GGAGCGTGTT GGCCAACGTT TTGTTCTGCC
    401  TTGCCGTGAT TTTTATCGGC ATCAGCGGCT GCGTGATGTG GTGGAAACGC
    451  CGTCCGTCCG GCGCGGTGGG CATGGTTCCG CCGGCGCAAA AAATCAAGCT
    501  GCCCGTCTGG TGGGCAATGG CGGTGCCGCT GCTGCTGATT GCATTGCTTT
    551  TCCCGACCGC GTTGCTTGCC ATTGCCGTGA TTTGGCTGTT GGATACGCTG
    601  CTGTTGTCGC GGATTCCTGT TTTGAGGAGA TGGTTTAAAT GA
```

This corresponds to the amino acid sequence <SEQ ID 780; ORF 219.a>:

```
a219.pep
      1  MTARLRKCRG FWSLRLCLFQ GRLWAKTVLT LTSR*HWKPS TVLRXEIGFK
     51  GRYQLNLPKG EDGVWTLSQD SMSYDMISPF ADRTVHIDQY SGKILADIRF
    101  DDYNPFGKFM AASIALHMGT LGWWSVLANV LFCLAVIFIG ISGCVMWWKR
    151  RPSGAVGMVP PAQKIKLPVW WAMAVPLLLI ALLFPTALLA IAVIWLLDTL
    201  LLSRIPVLRR WFK*
```

m219/a219 94.8% identity in 213 aa overlap

```
                        10        20        30        40        50        60
m219.pep        MTARLRKCRGFWSLRLCLFQGRXWAKTALTLTSRXHWKPSTALRGEIGFKGRYQLNLPKG
                |||||||||||||||||||||| ||||:|||||||||||||:|| |||||||||||||||
a219            MTARLRKCRGFWSLRLCLFQGRLWAKTVLTLTSRXHWKPSTVLRXEIGFKGRYQLNLPKG
                        10        20        30        40        50        60

                        70        80        90       100       110       120
m219.pep        EDGVWTLSQDSMSYDMISPFADRTVHIDQYSGKILADIRFDDYNPFGKFMAASIALHMGT
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a219            EDGVWTLSQDSMSYDMISPFADRTVHIDQYSGKILADIRFDDYNPFGKFMAASIALHMGT
                        70        80        90       100       110       120

                       130       140       150       160       170       180
m219.pep        LGWWSVLANVLFCLAVIFIGISGCVMWWKRRPTGAVGIVPPAQKVKLPVWWMMALPLLAI
                ||||||||||||||||||||||||||||||||:||||:||||||:|||||| ||:||| |
a219            LGWWSVLANVLFCLAVIFIGISGCVMWWKRRPSGAVGMVPPAQKIKLPVWWAMAVPLLLI
                       130       140       150       160       170       180

                       190       200       210
m219.pep        ALLFPTSLLAIAVIWLLDTLLLSRIPVLRRWFKX
                ||||||:|||||||||||||||||||||||||||
a219            ALLFPTALLAIAVIWLLDTLLLSRIPVLRRWFKX
                       190       200       210
```

```
m221.pep    CPAEVQLGKLVPSVQMWSEQFHFFFKIFDVGIGAAFVFGDDFVAAAVVADGVAERNVNVK
              |||||||||||||||||||||||||||||||||||||||;||||||||||:|||||
g221        AAAEVQLGKLVPSVQMWSEQFHFFFKIFDVGIGAAFVFGDDFVAAAVVADGVAKRNVNVK
                    130       140       150       160       170       180

m221.pep    GKRFVX
            ||||||
g221        GKRFVX
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 785>:

```
a221.seq
      1  ATGGTGGTTT TGATGCTCCG AAGTCTGGTG CGGCAGGCCG TAAATCAAAT
     51  CGACGCTGAC GGATTTGAAC CCCGCTTCGC GCGCCGCATC GATGACTTCT
    101  TTGGTTTCTT CGTAACTTTG GATGCGGTTG ACCGCCGCCT GCACTTTGGG
    151  GTCGAAATCC TGAATGCCGA TGCTCATGCG GTTGAAGCCG AGTCTGCCGA
    201  GCATGAGGAC GGTGTCGCGG CTGACTTTGC GCGGGTCGAT TTCGATGGAG
    251  TATTCGCCGG TGGGGATTAA CTCGAAATGT TTGCGTATCA TGCGGAAGAC
    301  ACGTTCGATT TGGTCGTCGC TCAAAAAGGT CGGCGTGCCG CCGCCGAAGT
    351  GCAGTTGGGC AAGCTGGTGC CGTCCGTTCA GATGTGGAGC GAGCAGTTCC
    401  ATTTCTTTTT CAAGAAATTC GATGTAGGCA TCGGCGCGGC TTTTGTCTTT
    451  GGTGATGATT TTGTTGCAGC CGCAGTAGTA GCAGATGGTG TTGCAGAACG
    501  GAATGTGAAT GTAAAGGGAA AGCGGTTTGT TTAA
```

This corresponds to the amino acid sequence <SEQ ID 786; ORF 221.a>:

```
a221.pep
      1  MVVLMLRSLV RQAVNQIDAD GFEPRFARRI DDFFGFFVTL DAVDRRLHFG
     51  VEILNADAHA VEAESAEHED GVAADFARVD FDGVFAGGD* LEMFAYHAED
    101  TFDLVVAQKG RRAAAEVQLG KLVPSVQMWS EQFHFFFKKF DVGIGAAFVF
    151  GDDFVAAAVV ADGVAERNVN VKGKRFV*
```

m221/a221 95.5% identity in 177 aa overlap

```
                    10        20        30        40        50        60
m221.pep    MXVLMXRSLVRQAVNQIDADGFEPRFARRIDDFFGFFVTLDAVDRRLHFGVEILNADAHA
            | ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
a221        MVVLMLRSLVRQAVNQIDADGFEPRFARRIDDFFGFFVTLDAVDRRLHFGVEILNADAHA
                    10        20        30        40        50        60

                    70        80        90       100       110       119
m221.pep    VEAESAEHEDGVAADFARVDFDGVFAGGDXLEMFAYHAEDTFDLFVAQKGA-CPAEVQLG
            |||||||||||||||||||||||||||||||||||||||||||| |||||   ||||||
a221        VEAESAEHEDGVAADFARVDFDGVFAGGDXLEMFAYHAEDTFDLVVAQKGRRAAAEVQLG
                    70        80        90       100       110       120

         120       130       140       150       160       170
m221.pep    KLVPSVQMWSEQFHFFFKIFDVGIGAAFVFGDDFVAAAVVADGVAERNVNVKGKRFVX
            |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
a221        KLVPSVQMWSEQFHFFFKKFDVGIGAAFVFGDDFVAAAVVADGVAERNVNVKGKRFVX
                    130       140       150       160       170
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 787>:

```
g223.seq
      1  atggaattca ggcaccaggt agtggtagtt ggtgtcgaac catttggtca
     51  tttcgatggc gaattggtct ttgttgccgc gcgccagttg gaagaattgt
    101  tccaaaggca ggttttggct atcgaagccg aaacgggcgg gaatcgcgcc
    151  cgtggatact tgcaggtcga ggatgtgatg gtagaaagtg aaatcacgta
    201  cagcaacgta atcagcgtta ggagcagctt ggtgtttcca gtttttctcg
    251  cgcaggtctt tggcaacgtc gagcagctct tgttcactga tctctttgcg
    301  ccagtatttt tcttgggcga atttcaattc acggaaggcg ccgacacgcg
    351  ggaagcctga
```

This corresponds to the amino acid sequence <SEQ ID 788; ORF 223.ng>:

g223.pep..

```
           70        80        90       100       110       120
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 793>:

```
g225.seq
     1  atggattctt ttttcaaacc ggcagtttgg gcggttttgt ggctgatgtt
    51  tgccgtccgc cccgcccttg ccgacgagtt gaccaacctg ctcagcagcc
   101  gcgagcagat tctcagacag tttgccgaag acgaacagcc cgttttaccc
   151  gtcaaccgag cccccgcccg gcgggcgggc aatgccgacg aactcatcgg
   201  cggcgcgatg gggcttaacg aacagcccgt tgtacgcgtc aaccgagccn
   251  ccgcccggcg ggcgggcaat gccgacaaac tcatcggcag cgcgatgcgg
   301  cttttgggta ttgcctaccg ctacggcggc acatcggtgt ctaccggttt
   351  tgactgcagc ggattcatgc agcacatctt caaacgcgcc atgggcatca
   401  acctgccgcg cacgtcggcg gaacaggcgc ggatgggcgc acccgttgcc
   451  cgaagcgaat tgcagcccgg ggatatggtg tttttccgca cgctcggcgg
   501  cagccgcatt tcccatgtcg gactttatat cggcaacaac cgcttcatcc
   551  acgcgccgcg cacggggaaa aatatcgaaa tcaccagcct gagccacaaa
   601  tattggagcg gcaaatatgc gttcgcccgc cgggtcaaga aaaacgaccc
   651  gtcacgcttt ctgaactga
```

This corresponds to the amino acid sequence <SEQ ID 794; ORF 225.ng>:

```
g225.pep
     1  MDSFFKPAVW AVLWLMFAVR PALADELTNL LSSREQILRQ FAEDEQPVLP
    51  VNRAPARRAG NADELIGGAM GLNEQPVVRV NRAXARRAGN ADKLIGSAMR
   101  LLGIAYRYGG TSVSTGFDCS GFMQHIFKRA MGINLPRTSA EQARMGAPVA
   151  RSELQPGDMV FFRTLGGSRI SHVGLYIGNN RFIHAPRTGK NIEITSLSHK
   201  YWSGKYAFAR RVKKNDPSRF LN*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 795>:

```
m225.seq  (partial)
     1  ..TTTTCAAACC CGGCAGTTTG GGCGGTTTTG TGGCTGAwGT TTGCCGTCCG
    51    CCCCGCCCTT GCCGACGAGT TGACCAACCT GCTCAGCAGC CGCGAGCAGA
   101    TTCTCAGACA GTTTGCCGAA GACGAACAGC CCGTTTTACC CATCAACCGA
   151    GCCCCCGCCC GGCGGGCGGG CAATGCCGAC GAACTCATCG GCAGCGCGAT
   201    GGGGCTTAAC GAACAGCCCG TTTTACCCGT CAACCGAGTC CCCGCCCGGC
   251    GGGCGGGCAA TGCCGACGAA CTCATCGGCA ACGCGATGGG GCTTAACGAA
   301    CAGCCCGTTT TACCCGTCAA CCGAGCCCCC GCCCGGCGGG CGGGCAATGC
   351    CGACGAACTC ATCGGCAACG CGATGGGACT TTTGGGTATT GCCTACCGCT
   401    ACGGCGGCAC ATCGGTTTCT ACCGGTTTTG ACTGCAGCGG CTTCATGCAG
   451    CACATCTTCA AACGCGCCAT GGGCATCAAC CTGCCGCGCA CGTCGGCAGA
   501    ACAGGCACGG ATGGGTACGC CGGTTGCCCG AAGCGAATTG CAGCCCGGAG
   551    ATATGGTGTT TTTCCGCACG CTCGGCGGCA GCCGCATTTC CCATGTCGGA
   601    CTTTATATCG GCAACAACCG CTTCATCCAC GCGCCGCGCA CGGGGAAAAA
   651    TATCGAAATC ACCAGCCTGA GCCACAAATA TTGGAGCGGC AAATACGCGT
   701    TCGCCCGCCG GGTCAAGAAA AACGACCCGT CCCGCTTTCT GAACTGA
```

This corresponds to the amino acid sequence <SEQ ID 796; ORF 225>:

```
m225.pep  (partial)
     1  ..FSNPAVWAVL WLXFAVRPAL ADELTNLLSS REQILRQFAE DEQPVLPINR
    51    APARRAGNAD ELIGSAMGLN EQPVLPVNRV PARRAGNADE LIGNAMGLNE
   101    QPVLPVNRAP ARRAGNADEL IGNAMGLLGI AYRYGGTSVS TGFDCSGFMQ
   151    HIFKRAMGIN LPRTSAEQAR MGTPVARSEL QPGDMVFFRT LGGSRISHVG
   201    LYIGNNRFIH APRTGKNIEI TSLSHKYWSG KYAFARRVKK NDPSRFLN*
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

ORF 225 shows 83.5% identity over a 248 aa overlap with a predicted ORF (ORF 225.ng) from *N. gonorrhoeae:*

```
m225/g225

                           10        20        30        40        50
m225.pep              FSNPAVWAVLWLXFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
                      | :||||||||| ||||||||||||||||||||||||||||||||||:|||||||||
g225               MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPVNRAPARRAG
```

```
                    ||||||||||||||||||||||                             |||||||||
a225                NADELIGSAMGLNEQPVLPVNRXPARRAGNADXLIGNAMGLNEQPVLPVNRVPARRAGNA
                            70        80        90       100       110       120

                    90        100       110       120       130       140
m225.pep            DELIGNAMGLNEQPVLPVNRAPARRAGNADELIGNAMGLLGIAYRYGGTSVSTGFDCSGF
                    ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
a225                DELIGNAMGLNEQPVLPVNRAPARRAGNADELIGNAMGLLGIAYRYGGTSISTGFDCSGF
                           130       140       150       160       170       180

                   150       160       170       180       190       200
m225.pep            MQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNNRF
                    |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
a225                MQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVXFRTLGGSRISHVGLYIGNNRF
                           190       200       210       220       230       240

                   210       220       230       240       249
m225.pep            IHAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDPSRFLNX
                    |||||||||||||||||||||||||||||||||||||||||
a225                IHAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDPSRFLNX
                           250       260       270       280
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 799>:

g225-1.seq

```
  1  atggattctt ttttcaaacc ggcagtttgg gcggttttgt ggctgatgtt
 51  tgccgtccgc cccgcccttg ccgacgagtt gaccaacctg ctcagcagcc
101  gcgagcagat tctcagacag tttgccgaag acgaacagcc cgttttaccc
151  gtcaaccgag cccccgcccg gcgggcgggc aatgccgacg aactcatcgg
201  cggcgcgatg gggcttaacg aacagcccgt tgtacgcgtc aaccgagccn
251  ccgcccggcg ggcgggcaat gccgacaaac tcatcggcag cgcgatgcgg
301  cttttgggta ttgcctaccg ctacggcggc acatcggtgt ctaccggttt
351  tgactgcagc ggattcatgc agcacatctt caaacgcgcc atgggcatca
401  acctgccgcg cacgtcggcg gaacaggcgc ggatgggcgc acccgttgcc
451  cgaagcgaat tgcagcccgg ggatatggtg tttttccgca cgctcggcgg
501  cagccgcatt tcccatgtcg gactttatat cggcaacaac cgcttcatcc
551  acgcgccgcg cacggggaaa aatatcgaaa tcaccagcct gagccacaaa
601  tattggagcg gcaaatatgc gttcgcccgc cgggtcaaga aaaacgaccc
651  gtcacgcttt ctgaactga
```

This corresponds to the amino acid sequence <SEQ ID 800; ORF 225-1.ng>:

g225-1.pep

```
  1  MDSFFKPAVW AVLWLMFAVR PALADELTNL LSSREQILRQ FAEDEQPVLP
 51  VNRAPARRAG NADELIGGAM GLNEQPVVRV NRAXARRAGN ADKLIGSAMR
101  LLGIAYRYGG TSVSTGFDCS GFMQHIFKRA MGINLPRTSA EQARMGAPVA
151  RSELQPGDMV FFRTLGGSRI SHVGLYIGNN RFIHAPRTGK NIEITSLSHK
201  YWSGKYAFAR RVKKNDPSRF LN*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 801>:

m225-1.seq

```
  1  ATGGATTCTT TTTTCAAACC GGCAGTTTGG GCGGTTTTGT GGCTGATGTT
 51  TGCCGTCCGC CCCGCCCTTG CCGACGAGTT GACCAACyTG CTCAGCAGCC
101  GCGAGCAGAT TCTCAGACAG TTTGCCGAAG ACGAACAGCC CGTTTTACCC
151  ATCAACCGAG CCCCCGCCCG GCGGGCGGGC AATGCCGACG AACTCATCGG
201  CAGCGCGATG GGGCTTAACG AACAGCCCGT TTTACCCGTC AACCGAGTCC
251  CCGCCCGGCG GGCGGGCAAT GCCGACGAAC TCATCGGCAA CGCGATGGGG
301  CTTAACGAAC AGCCCGTTTT ACCCGTCAAC CGAGCCCCCG CCCGGCGGGC
351  GGGCAATGCC GACGAACTCA TCGGCAACGC GATGGGACTT TTGGGTATTG
401  CCTACCGCTA CGGCGGCACA TCGGTTTCTA CCGGTTTTGA CTGCAGCGGC
451  TTCATGCAGC ACATCTTCAA ACGCGCCATG GGCATCAACC TGCCGCGCAC
501  GTCGGCAGAA CAGGCACGGA TGGGTACGCC GGTTGCCCGA AGCGAATTGC
551  AGCCCGGAGA TATGGTGTTT TTCCGCACGC TCGGCGGCAG CCGCATTTCC
601  CATGTCGGAC TTTATATCGG CAACAACCGC TTCATCCACG CGCCGCGCAC
651  GGGGAAAAAT ATCGAAATCA CCAGCCTGAG CCACAAATAT TGGAGCGGCA
701  AATACGCGTT CGCCCGCCGG GTCAAGAAAA ACGACCCGTC CCGCTTTCTG
751  AACTGA
```

```
a225-1/m225-1     88.6% identity in 280 aa overlap

                   10        20        30        40        50        60
a225-1.pep  MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRXPARRAG
            |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
m225-1      MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
                   10        20        30        40        50        60

                   70        80        90       100       110       120
a225-1.pep  NADELIGSAMGLNEQPVLPVNRXPARRAGNADXLIGNAMGLNEQPVLPVNRVPARRAGNA
            |||||||||||||||                            |||||||||||||||
m225-1      NADELIGSAMGLNEQP---------------------------VLPVNRVPARRAGNA
                   70                                        80        90

                  130       140       150       160       170       180
a225-1.pep  DELIGNAMGLNEQPVLPVNRAPARRAGNADELIGNAMGLLGIAYRYGGTSISTGFDCSGF
            ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
m225-1      DELIGNAMGLNEQPVLPVNRAPARRAGNADELIGNAMGLLGIAYRYGGTSVSTGFDCSGF
                 100       110       120       130       140       150

                  190       200       210       220       230       240
a225-1.pep  MQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVXFRTLGGSRISHVGLYIGNNRF
            |||||||||||||||||||||||:|||||||||||||| |||..|||||||||||||||
m225-1      MQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNNRF
                 160       170       180       190       200       210

                  250       260       270       280
a225-1.pep  IHAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDPSRFLNX
            |||||||||||||||||||||||||||||||||||||||||
m225-1      IHAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDPSRFLNX
                 220       230       240       250
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 805>:

```
g226.seq
  1   ATGAGCGAAA TCCTCAGGCA GCCCAGCGTT CTGCTTTTCC TCACGCTTGC
 51   CGTGTACGCG CTTGCGATTA TCGTGCGCAC GCGCACGGGC AATATCTTCT
101   GCAACCCCGT ACTCGTCAGC ACTATCGTGC TGATTGCCTA CCTGAAAATC
151   CTCGGTATCG ATTATGCGGT GTACCACAAC GCCGCGCAAT TCATTGATTT
201   TCGGCTGAAA cccGccgtCG TCGTGCTTGC CGTGCCGCTC TACCAAAACC
251   GCCGTAAAAT CTTCAACCAG TGGCTGCCCG TCATCGTTTC GCAGCTTGCG
301   GGCAGCGTTA cggGCATTGT tacggggATG TATTTTgccg cttggctcgg
351   gccggatacc caattctcct tcccgcctcg tcttcaatat ctgttattta
401   cacccctctgg aatcccaatt cacaccctgt atgcgcgggt tctcccgcca
451   tttctgttgc ctccgcctct cctgccgcgc ctcggcccgc atacattgcg
501   ccggttcaca atacttccaa aaaaactacg gccgtttaag cccctcctcc
551   cagttgtggt cctttctcct Ccgggcctcg cccctcccct cttataa
```

This corresponds to the amino acid sequence <SEQ ID 806; ORF 226.ng>:

```
g226.pep
  1   MSEILRQPSV LLFLTLAVYA LAIIVRTRTG NIFCNPVLVS TIVLIAYLKI
 51   LGIDYAVYHN AAQFIDFRLK PAVVVLAVPL YQNRRKIFNQ WLPVIVSQLA
101   GSVTGIVTGM YFAAWLGPDT QFSFPPRLQY LLFTPSGIPI HTLYARVLPP
151   FLLPPPLLPR LGPHTLRRFT ILPKKLRPFK PLLPVVVLSP PGLAPPLL*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 807>:

```
m226.seq
  1   ATGAACGAAA TCCTCAGGCA GCCCAGCGTT CTGCTTTTCC TCACGCTTGC
 51   CGTGTACGCG CTTGCGATTA TCGtGCGCAC GCGCACGGGC AATATCTTCT
101   GCAACCCCGT ACTCGTCAGC ACTATCGTGC TGATTGCCTA CCTGAAAATC
151   CTCGGTATCG ATTATGCGGT GTACCACAAC GCCGCGCAAT TCATTGATTT
201   TTGGCTGAAA CCCGCCGTCG TCGTGCTTGC CGTGCCGCTC TACCAAAACC
251   GCCGTAAAAT CTTCAACCAG TGGCTGCCCG TCATCGTTTC ACAGCTTGCG
301   GGCAGCGTTA CGGGCATTGT TACAGGGATG TATTTTGCCA AATGGCTGGG
351   CGCGGAACGC GAAGTCGTCC TCTCGCTCGC GTCCAAATCT GTTACCAACC
401   CCATCGCTAT TGAAATCACC CGCTCCATCG GCGGCATTCC CGCCATTACC
451   GCCGCCACCG TCATCATTGC CGGTCTGGTC GGACAGATTG CCGGTTACAA
```

```
m226.pep    MNEILRQPSVLLFLTLAVYALAIIVRTRTGNIFCNPVLVSTIVLIAYLKILGIDYAVYHN
            ||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
a226        MNEILRQPSILLFLTLAVYALAIIVRTRTGNIFCNPVLVSTIVLIAYLKILGIDYAVYHN
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m226.pep    AAQFIDFWLKPAVVVLAVPLYQNRRKIFNQWLPVIVSQLAGSVTGIVTGMYFAKWLGAER
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a226        AAQFIDFWLKPAVVVLAVPLYQNRRKIFNQWLPVIVSQLAGSVTGIVTGMYFAKWLGAER
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m226.pep    EVVLSLASKSVTNPIAIEITRSIGGIPAITAATVIIAGLVGQIAGYKMLKNTVVMPSSVG
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a226        EVVLSLASKSVTNPIAIEITRSIGGIPAITAATVIIAGLVGQIAGYKMLKNTVVMPSSVG
                   130       140       150       160       170       180

                   190       200       210       220       230
m226.pep    MSLGTASHAMGIAASLERSRRMAAYAGLGLTFNGVLTALIAPLLIPVLGFX
            |||||||||||||||||||||||||||||||||||||||||||||||||||
a226        MSLGTASHAMGIAASLERSRRMAAYAGLGLTFNGVLTALIAPLLIPVLGFX
                   190       200       210       220       230
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 811>:

```
g227.seq
      1  atgaacatca tccgcgcgct cctcatcatc ctcggctgcc tcgccgccgg
     51  cgaaaccgcc gttttcctag caggcatcaa actgcccggc agcatcgtcg
    101  gcatgggcgt gctgtttgcg cttttgcagg cgggttggct caaaacgtct
    151  tggctgcaac agcttaccga cgcgctgatg gcaaacctga cgctgttcct
    201  cgtgccgccc tgcgtggcgg tcatcagcta tttggatttg attgccgacg
    251  attggttttc gatactggtt tccgcctccg ccagcacttt gtgcgtactg
    301  ctggttacgg gcaaggttca ccgctggata cggagcatta tctga
```

This corresponds to the amino acid sequence <SEQ ID 812; ORF 227.ng>:

```
g227.pep
      1  MNIIRALLII LGCLAAGETA VFLAGIKLPG SIVGMGVLFA LLQAGWLKTS
     51  WLQQLTDALM ANLTLFLVPP CVAVISYLDL IADDWFSILV SASASTLCVL
    101  LVTGKVHRWI RSII*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 813>:

```
m227.seq  (partial)
      1  ..ACGTCTTkGC TGCAACAGCT TACCGACGCG CTGATGTCGA ACCTGACGCT
     51    GTtCCTCGTG CCgCC.TGCG TGGCGGTCAT CAGCTATTTG GATTTGATTG
    101    CCGACGATTG GTTTTCGATA CTGGTTTCCG CCTCCGCCAG cACTTTGTGC
    151    GTACTGCTGG TTACGGGCAA AGTCCACCGG TGGATACGGG GTATTATCCG
    201    ATGA
```

This corresponds to the amino acid sequence <SEQ ID 814; ORF 227>:

```
m227.pep  (partial)
      1  ..TSXLQQLTDA LMSNLTLFLV PPCVAVISYL DLIADDWFSI LVSASASTLC
     51    VLLVTGKVHR WIRGIIR*
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

ORF 227 shows 95.5% identity over a 66 aa overlap with a predicted ORF (ORF 227.ng) from *N. gonorrhoeae:*

```
m227/g227

                                                10        20        30
m227.pep                                TSXLQQLTDALMSNLTLFLVPPCVAVISYL
                                        || |||||||||:|||||||||||||||||
g227        TAVFLAGIKLPGSIVGMGVLFALLQAGWLKTSWLQQLTDALMANLTLFLVPPCVAVISYL
              20        30        40        50        60        70

                    40        50        60
```

```
a228.pep
      1  MKKLLIAAMM AAALAACSQE AKQEVKEAVQ AVESDVKDTA ASAAESAASA
     51  VEEAKDQVKD AAADAKASAE EAVTEAKEAV TEAAKDTLNK AADATQEAAD
    101  KMKDAAK*
```

m228/a228 100.0% identity in 107 aa overlap

```
                    10        20        30        40        50        60
m228.pep    MKKLLIAAMMAAALAACSQEAKQEVKEAVQAVESDVKDTAASAAESAASAVEEAKDQVKD
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a228        MKKLLIAAMMAAALAACSQEAKQEVKEAVQAVESDVKDTAASAAESAASAVEEAKDQVKD
                    10        20        30        40        50        60

                    70        80        90       100
m228.pep    AAADAKASAEEAVTEAKEAVTEAAKDTLNKAADATQEAADKMKDAAKX
            ||||||||||||||||||||||||||||||||||||||||||||||||
a228        AAADAKASAEEAVTEAKEAVTEAAKDTLNKAADATQEAADKMKDAAKX
                    70        80        90       100
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 821>:

```
g229.seq
      1  atggctgccg tatcgggcgg cggtgcggtc ttcctgataa tgcttccaca
     51  tattgcccgc gttcagcgtc agccgccagc gttcgcccaa gcgtcgggag
    101  aaatcggcat tgaagccgcc ggcgaaattg tatcggctgc cgcccaagag
    151  gttttgcccg acaaacggca cggtgccgaa cgagcgcgtt accgaacggt
    201  tttgatggcc gaacgacagg cgcaggttct gttcgctgaa atctttgtta
    251  tcccaataat gcacgccgcg gctgatgccg ccgtagagga aatgatgccc
    301  gcccgcattg atttcgcgcg acacgcccaa gccgtagcgc aaaccgtgtg
    351  ccttttgcgg caggctgtcg gcggttttcg tccagcttct gcccgcaaat
    401  tcaatcgttt tttcggacga agcgttgttt atagcggatt aacaaaaatc
    451  aggacaaggc ggcgggccgc aggcagtacg gatggtacgg aaccggttcg
    501  cccggtgctt ggacgcctta gggaaccgtt ccctttgagc cggggcgggg
    551  caacccgtac cggtttttgt tcatccgcca tattgtgttg a
```

This corresponds to the amino acid sequence <SEQ ID 822; ORF 229.ng>:

```
g229.pep
      1  MAAVSGGGAV FLIMLPHIAR VQRQPPAFAQ ASGEIGIEAA GEIVSAAAQE
     51  VLPDKRHGAE RARYRTVLMA ERQAQVLFAE IFVIPIMHAA ADAAVEEMMP
    101  ARIDFARHAQ AVAQTVCLLR QAVGGFRPAS ARKFNRFFGR SVVYSGLTKI
    151  RTRRRAAGST DGTEPVRPVL GRLREPFPLS RGGATRTGFC SSAILC*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 823>:

```
m229.seq (partial)
      1  ..GCTCAAGCGT TGGGAGAAAT CGGCATTGAA GCCGCCGACG AAATTGTATC
     51    GGCTGCCGCC TAAGAGGTTT TGCTCGACAA ACGGCACGAT GCCGAACGAG
    101    CGCGTTACCG AACGGTTTTT ATAGCCGAAC GACAGGCGCA GGCTCTGTTC
    151    GCTGAAATCT TTGTTATCCC AATAATGCAC GCCGCCGCCG CTGATGCCGC
    201    CGTAGAGGAA ATGATGCCTG CCCGCATTGA TTTCGCGCGA CACGCCTAAG
    251    CCCTAGCGCA AACCGTGTGC CTTTTGCGGC AGGCTGTCGG CGGTTTTCGT
    301    CCAGCTTCTG CCCGCAAATT CAATCGTTTT TTCGGACGAA GCGTTGTTTA
    351    TAGCGGATTA ACAAAAATCA GGACAAGGCA ACGAAGCCGC AGACAGTACA
    401    AATAGTACGG AACCGATTCA CTTGGTGCTT CAGCACcTTA GAGAATCGTT
    451    CTCTTTTTTG TTCATCCGCT ATATTGTGTT GA
```

This corresponds to the amino acid sequence <SEQ ID 824; ORF 229>:

```
m229.pep  (partial)
      1  ..AQALGEIGIE AADEIVSAAA XEVLLDKRHD AERARYRTVF IAERQAQALF
     51    AEIFVIPIMH AAAADAAVEE MMPARIDFAR HAXALAQTVC LLRQAVGGFR
    101    PASARKFNRF FGRSVVYSGL TKIRTRQRSA DSTNSTEPIH LVLQHLRESR
    151    SLFCSSAILC *
```

Computer analysis of this amino acid sequence gave the following results:

```
                100         110         120         130         140       149
m229.pep      RQAVGGFRPASARKFNRFFGRSVVYSGLTKIRTRQRSADSTNSTEPIHLVLQHLRES---
              |||||||||||| ||||||||||||||||||||||:|||||:||||||||||||||
a229          RQAVGGFRPASACKFNRFFGRSVVYSGLTKIRTRRRSADSTDSTEPIHLVLQHLRESSLX
                  130         140         150         160         170        180

              150         160
m229.pep      ------RSLFCSSAILCX
                    |:  ||||:|
a229          AKARQRRTGFCSSTI
                  190
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 827>:

```
g230.seq
   1  atgttccatt ccatcgaaaa atacagaaca cccgcccaag tcttattagg
  51  cctgattgca ttaacttttg tcggcttcgg cgtcagcacg gtttcccatc
 101  cgggcgccga ctacatcgtc caagtgggcg acgaaaaaat cagcgagcac
 151  tcaaccaaca acgccatgca gaacgagcag gcggacggcg gcagcccttg
 201  gcgcgacgcg gtgttccaat ccctgctgca acgcgcctac ctgaaacagg
 251  gcgcgaagct gatgggcatt tcggtttctt ccgaacaaat caagcagatg
 301  attgtggacg atcccaattt ccacgacgca aacggcaaat tcagtcacgc
 351  gcttttgagt caatacctgt cgcaacgcca tatgtctgaa gaccagtttg
 401  tcgaagaaat ccgcgatcag tttgccttgc agaatttggt aagcctcgtc
 451  caaaacggcg tattggtcgg cgacgcgcag gcggaacagc tgatcaggct
 501  gacgcaggtc aaccgcacca tccgttcgca cactttcaac cccgacgagt
 551  tcatcgccca agtcaaagcg tctgaagccg atttgcagaa attttataat
 601  gcgaacaaaa aagactatct gctgccgcag gcggtcaaat tggaatatgt
 651  cgccttgaat ctgaaggatt ttgcagacaa gcagaccgtc agtgaaacgg
 701  aagtgaaaaa tgcgtttgaa gagcgcgtgg cgcgtttgcc ggcacatgaa
 751  gccaaacctt ctttcgagca ggaaaaagcc gccgtcgaaa acgaattgaa
 801  aatgaaaaag gcggttgccg acttcaacaa ggcaaaagaa aagctgggcg
 851  acgatgcgtt caatcatccc tcctcgcttg ccgaagccgc caaaaacagc
 901  ggtttgaaag tggaaaccca agaaacttgg ctgagcaggc aggacgcaca
 951  aatgtccggc atgcccgaaa acctaatcaa tgccgtattc agcgacgacg
1001  tattgaagaa aaaacacaat tccgaagtgc tgaccatcaa cagcgaaacc
1051  gcgtgggtcg tccgcgccaa agaagtccgc gaagaaaaaa acctactgtt
1101  tgaagaagcc aaagatgcgg tgcgtcaggc ctatatccgt accgaagccg
1151  ccaaactttt gaaaacaatg taa
```

This corresponds to the amino acid sequence <SEQ ID 828; ORF 230.ng>:

```
g230.pep
   1  MFHSIEKYRT PAQVLLGLIA LTFVGFGVST VSHPGADYIV QVGDEKISEH
  51  SINNAMQNEQ ADGGSPWRDA VFQSLLQRAY LKQGAKLMGI SVSSEQIKQM
 101  IVDDPNFHDA NGKFSHALLS QYLSQRHMSE DQFVEEIRDQ FALQNLVSLV
 151  QNGVLVGDAQ AEQLIRLTQV NRTIRSHTFN PDEFIAQVKA SEADLQKFYN
 201  ANKKDYLLPQ AVKLEYVALN LKDFADKQTV SETEVKNAFE ERVARLPAHE
 251  AKPSFEQEKA AVENELKMKK AVADFNKAKE KLGDDAFNHP SSLAEAAKNS
 301  GLKVETQETW LSRQDAQMSG MPENLINAVF SDDVLKKKHN SEVLTINSET
 351  AWVVRAKEVR EEKNLLFEEA KDAVRQAYIR TEAAKLLKTM *
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 829>:

```
m230.seq (partial)
   1  ATGTTCCATT CCATCGAAAA ATACAGAACG CCCGCCCAAG TCCTTTTGGG
  51  CCTGATTGCA TTAACCTTCG TCGGCTTCGG GGTCAGCACG GTATCCCATC
 101  CGGGTGCCGA CTACATCGTC CAAGTGGGCG ACGAAAAAaT CAGCGACCAC
 151  TCCATCAACA ACGCCATACA GAACGAACAG GCGGACGGCG GCGGCCCTTC
 201  GCc.GACGCG GTGTTCCAAT CCCTGCTGCA ACGCGCCTAC CTGAAACAGG
 251  GCGCGAAGCT GATGGGCATT TCGGTTTCTT CCGAACAAAT CAAGCAAATT
 301  ATCGTGGACG ATCCCAATTT CCACGACGCA AACGGCAAAT TCGACCACGC
 351  GCTTTTAAAC CGCTACCTTT CCCAACGCCA TATGTCTGAA GACCAGTTTG
 401  TCGAAGAAAT CCGCGATCAG TTTGCCTTGC AGAATTTGGT AAACCTCGTC
```

```
                    310       320       330       340       350       360
m230.pep    GLKVETQETWLSRQDAQMSGMPENLINAVFSDDVLKKKHNSEVLTINSETAWVVRAKEVR
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g230        GLKVETQETWLSRQDAQMSGMPENLINAVFSDDVLKKKHNSEVLTINSETAWVVRAKEVR
                    310       320       330       340       350       360

                    370       380
m230.pep    EEKTLPFAEAKDAVRQAYIRTEAAKL
            |||:| | ||||||||||||||||||
g230        EEKNLLFEEAKDAVRQAYIRTEAAKLLKTM
                    370       380       390
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 831>:

```
a230.seq  (partial)
       1  ATGTTCCATT CCATCGAAAA ATACAGAACG CCCGCCCAAG TCCTTTTGGG
      51  CCTGATTGCA TTAACCTTCG TCGGCTTCGG GGTCAGCACG GTATCCCATC
     101  CGGGTGCCGA CTACATCGTC CAAGTGGGCG ACGAAAAAAT CAGCGACCAC
     151  TCCATCAACA ACGCCATACA GAACGAACAG GCGGACGGCG GCGGCCCTTC
     201  GCGCGACGCG GTGTTCCAAT CCCTGCTACA ACGCGCCTAC CTGAAACAGG
     251  GCGCGAAGCT GATGGGCATT TCGGTTTCTT CCGAACAAAT CAAGCAGATT
     301  ATCGTGGACG ATCCCAATTT CCACGACGCA AACGGCAAAT TCGACCACGC
     351  GCTTTTAAAC CGCTACCTTT CCCAACGTCA TATGTCTGAA GACCAGTTTG
     401  TCGAAGAAAT CCGCGATCAG TTTGCCTTGC CGAATTTGGT AAACCTCGTC
     451  CAAAACGGCG TATTGGTCGG CGACGCGCAG GCGGAACAGC TGATCAGGCT
     501  GACGCAGGTC AACCGCACCA TCCGTTCGCA CACTTTCAAC CCCGACGAAT
     551  TCATCGCCCA AGTCAAAGTG TCTGAAGCCG ATTTGCAGAA GTTTTATAAC
     601  GCAAACAAAA AAGACTACCT GCTTCCCAAA GCGGTCAAAT TGGAATATGT
     651  CGCCTTGAAT CTGAAAGACT TTGCAGACAA ACAGACCGTC AGCGAAACAG
     701  AAGTGAAAAA TGCGTTTGAA GAGCGCGTGG CGCGTTTGCC GGCAAATGAA
     751  GCCAAACCTT CTTTCGAGCA GGAAAAAGCC GCCGTCGAAA ACGAATTGAA
     801  AATGAAAAAG GCGGTTGCCG ACTTCAATAA GGCAAAAGAA AAGCTGGGCG
     851  ATGACGCGTT CAACCATCCT TCCTCGCTTG CCGAAGCCGC CAAAAACAGC
     901  GGTTTGAAAG TCGAAACCCA AGAAACTTGG CTGAGCAGGC AGGATGCGCA
     951  AATGTCCGGT ATGCCCGAAA ACCTGATCAA TGCCGTATTC AGCGACGACG
    1001  TATTGAAGAA AAAACACAAT TCCGAAGTGC TGACCATCAA CAGCGAAACC
    1051  GCGTGGGTCG TCCGCGCCAA AGAAGTCCGC GAAGAGAAAA CCCTGCCGTT
    1101  TGCCGAAGCC AAAGACGCGG TACGTCAGGC TTATATCCGT ACCGAAGCCG
    1151  CCAAACTT
```

This corresponds to the amino acid sequence <SEQ ID 832; ORF 230.a>:

```
a230.pep  (partial)
       1  MFHSIEKYRT PAQVLLGLIA LTFVGFGVST VSHPGADYIV QVGDEKISDH
      51  SINNAIQNEQ ADGGGPSRDA VFQSLLQRAY LKQGAKLMGI SVSSEQIKQI
     101  IVDDPNFHDA NGKFDHALLN RYLSQRHMSE DQFVEEIRDQ FALQNLVNLV
     151  QNGVLVGDAQ AEQLIRLTQV NRTIRSHTFN PDEFIAQVKV SEADLQKFYN
     201  ANKKDYLLPK AVKLEYVALN LKDFADKQTV SETEVKNAFE ERVARLPANE
     251  AKPSFEQEKA AVENELKMKK AVADFNKAKE KLGDDAFNHP SSLAEAAKNS
     301  GLKVETQETW LSRQDAQMSG MPENLINAVF SDDVLKKKHN SEVLTINSET
     351  AWVVRAKEVR EEKTLPFAEA KDAVRQAYIR TEAAKL
```

m230/a230 99.2% identity in 386 aa overlap

```
                    10        20        30        40        50        60
m230.pep    MFHSIEKYRTPAQVLLGLIALTFVGFGVSTVSHPGADYIVQVGDEKISDHSINNAIQNEQ
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a230        MFHSIEKYRTPAQVLLGLIALTFVGFGVSTVSHPGADYIVQVGDEKISDHSINNAIQNEQ
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m230.pep    ADGGGPSPDAVFQSLLQRAYLKQGAKLMGISVSSEQIKQIIVDDPNFHDANGKFDHALLN
            ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
a230        ADGGGPSRDAVFQSLLQRAYLKQGAKLMGISVSSEQIKQIIVDDPNFHDANGKFDHALLN
                    70        80        90       100       110       120

                   130       140       150       160       170       180
```

```
151  QNGVLVGDAQ AEQLIRLTQV NRTIRSHTFN PDEFIAQVKA SEADLQKFYN
201  ANKKDYLLPQ AVKLEYVALN LKDFADKQTV SETEVKNAFE ERVARLPAHE
251  AKPSFEQEKA AVENELKMKK AVADFNKAKE KLGDDAFNHP SSLAEAAKNS
301  GLKVETQETW LSRQDAQMSG MPENLINAVF SDDVLKKKHN SEVLTINSET
351  AWVVRAKEVR EEKNLLFEEA KDAVRQAYIR TEAAKLAENK AKEVLTQLNG
401  GKAVDVKWSE VSVLGAQQAR QSMPPEAYAE LLKAKPANGK PAYVRLTGLP
451  APVIVEAQAV TPPEDIAAQL PPAKQALAQQ QSANTFDLLI RYFNGKIKQT
501  KGAQSVDNGD GQ*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 835>:

m230-1.seq

```
   1  ATGTTCCATT CCATCGAAAA ATACAGAACG CCCGCCCAAG TCCTTTTGGG
  51  CCTGATTGCA TTAACCTTCG TCGGCTTCGG GGTCAGCACG GTATCCCATC
 101  CGGGTGCCGA CTACATCGTC CAAGTGGGCG ACGAAAAAAT CAGCGACCAC
 151  TCCATCAACA ACGCCATACA GAACGAACAG GCGGACGGCG GCGGCCCTTC
 201  GCGCGACGCG GTGTTCCAAT CCCTGCTGCA ACGCGCCTAC CTGAAACAGG
 251  GCGCGAAGCT GATGGGCATT TCGGTTTCTT CCGAACAAAT CAAGCAAATT
 301  ATCGTGGACG ATCCCAATTT CCACGACGCA AACGGCAAAT TCGACCACGC
 351  GCTTTTAAAC CGCTACCTTT CCCAACGCCA TATGTCTGAA GACCAGTTTG
 401  TCGAAGAAAT CCGCGATCAG TTTGCCTTGC AGAATTTGGT AAACCTCGTC
 451  CAAAACGGCG TATTGGTCGG CGACGCGCAG GCGGAACAGC TGATCAGGCT
 501  GACACAGGTC AACCGCACCA TCCGTTCGCA CACTTTCAAC CCCGACGAGT
 551  TCATCGCCCA AGTCAAAGTG TCTGAAGCCG ATTTGCAGAA ATTTTATAAT
 601  GCGAACAAAA AAGACTATCT GCTGCCGCAG GCGGTCAAAT TGGAATATGT
 651  CGCCTTGAAT CTGAAGGATT TTGCAGACAA GCAGACCGTC AGTGAAACGG
 701  AAGTGAAAAA TGCATTTGAA GAGCGCGTGG CGCGTTTGCC GGCAAATGAA
 751  GCCAAACCTT CTTTCGAGCA GGAAAAAGCC GCCGTCGAAA ACGAATTGAA
 801  AATGAAAAAG GCGGTTGCCG ACTTCAACAA GGCAAAAGAA AAATTGGGCG
 851  ACGATGCGTT CAACCATCCT TCCTCGCTTG CCGAAGCCGC CAAAAACAGC
 901  GGTTTGAAAG TCGAAACCCA AGAAACTTGG CTGAGTAGGC AGGACGCGCA
 951  AATGTCCGGT ATGCCCGAAA ACCTGATCAA TGCCGTATTC AGCGACGACG
1001  TATTGAAGAA AAAACACAAT TCCGAAGTGC TGACCATCAA CAGCGAAACC
1051  GCGTGGGTCG TCCGCGCCAA AGAAGTCCGC GAAGAGAAAA CCCTGCCGTT
1101  TGCCGAAGCC AAAGACGCGG TACGTCAGGC TTATATCCGT ACCGAAGCCG
1151  CCAAACTTGC CGAAAACAAG GCAAAAGACG TGCTTACCCA ACTGAACGGC
1201  GGCAAGGCTG TTGACGTGAA ATGGTCGGAA GTGTCCGTTT TGGGCGCACA
1251  GCAGGCAAGG CAGTCCATGC CGCCCGAGGC TTATGCGGAA CTGCTGAAAG
1301  CAAAACCGGC AAACGGCAAA CCCGCCTACG TCAGGCTGAT CGGTCTGCCG
1351  GCACCCGTGA TTGTCGAAGT ACAGGCTGTA ACCCCGCCGG ATGATATCGC
1401  CGCACAGCTT CCGCTTGCAA AACAGGCTTT GGCGCAACAG CAGTCTGCCA
1451  ATACTTTCGA CTTGTTGATA CGTTATTTCA ACGGCAAAAT CAAACAGACC
1501  AAAGGAGCGC AATCGGTCGA CAACGGCGAC GGTCAGTAA
```

This corresponds to the amino acid sequence <SEQ ID 836; ORF 230-1>:

m230-1.pep

```
  1  MFHSIEKYRT PAQVLLGLIA LTFVGFGVST VSHPGADYIV QVGDEKISDH
 51  SINNAIQNEQ ADGGGPSRDA VFQSLLQRAY LKQGAKLMGI SVSSEQIKQI
101  IVDDPNFHDA NGKFDHALLN RYLSQRHMSE DQFVEEIRDQ FALQNLVNLV
151  QNGVLVGDAQ AEQLIRLTQV NRTIRSHTFN PDEFIAQVKV SEADLQKFYN
201  ANKKDYLLPQ AVKLEYVALN LKDFADKQTV SETEVKNAFE ERVARLPANE
251  AKPSFEQEKA AVENELKMKK AVADFNKAKE KLGDDAFNHP SSLAEAAKNS
301  GLKVETQETW LSRQDAQMSG MPENLINAVF SDDVLKKKHN SEVLTINSET
351  AWVVRAKEVR EEKTLPFAEA KDAVRQAYIR TEAAKLAENK AKDVLTQLNG
401  GKAVDVKWSE VSVLGAQQAR QSMPPEAYAE LLKAKPANGK PAYVRLIGLP
451  APVIVEVQAV TPPDDIAAQL PLAKQALAQQ QSANTFDLLI RYFNGKIKQT
501  KGAQSVDNGD GQ*
```

m230-1/g230-1 96.3% identity in 512 aa overlap

```
                    10        20        30        40        50        60
m230-1.pep  MFHSIEKYRTPAQVLLGLIALTFVGFGVSTVSHPGADYIVQVGDEKISDHSINNAIQNEQ
            ||||||||||||||||||||||||||||||||||||||||||||||||:||||||:||||
g230-1      MFHSIEKYRTPAQVLLGLIALTFVGFGVSTVSHPGADYIVQVGDEKISEHSINNAMQNEQ
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m230-1.pep  ADGGGPSRDAVFQSLLQRAYLKQGAKLMGISVSSEQIKQIIVDDPNFHDANGKFDHALLN
            ||||:| |||||||||||||||||||||||||||||||||:||||||||||||||:||||:
g230-1      ADGGSPWRDAVFQSLLQRAYLKQGAKLMGISVSSEQIKQMIVDDPNFHDANGKFSHALLS
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m230-1.pep  RYLSQRHMSEDQFVEEIRDQFALQNLVNLVQNGVLVGDAQAEQLIRLTQVNRTIRSHTFN
```

This corresponds to the amino acid sequence <SEQ ID 838; ORF 230-1.a>:

a230-1.pep

```
  1  MFHSIEKYRT PAQVLLGLIA LTFVGFGVST VSHPGADYIV QVGDEKISDH
 51  SINNAIQNEQ ADGGGPSRDA VFQSLLQRAY LKQGAKLMGI SVSSEQIKQI
101  IVDDPNFHDA NGKFDHALLN RYLSQRHMSE DQFVEEIRDQ FALQNLVNLV
151  QNGVLVGDAQ AEQLIRLTQV NRTIRSHTFN PDEFIAQVKV SEADLQKFYN
201  ANKKDYLLPK AVKLEYVALN LKDFADKQTV SETEVKNAFE ERVARLPANE
251  AKPSFEQEKA AVENELKMKK AVADFNKAKE KLGDDAFNHP SSLAEAAKNS
301  GLKVETQETW LSRQDAQMSG MPENLINAVF SDDVLKKKHN SEVLTINSET
351  AWVVRAKEVR EEKTLPFAEA KDAVRQAYIR TEAAKLAENK AKDVLTQLNG
401  GKAVDVKWSE VSVLGAQQAR QSMPPEAYAE LLKAKPANGK PAYVRLIGLP
451  APVIVEVQAV TPPDDIAAQL PLAKQALAQQ QSANTFDLLI RYFNGKIKQT
501      KGAQSVDNGD GQ*
```

a230-1/m230-1 99.8% identity in 512 aa overlap

```
                    10        20        30        40        50        60
a230-1.pep  MFHSIEKYRTPAQVLLGLIALTFVGFGVSTVSHPGADYIVQVGDEKISDHSINNAIQNEQ
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m230-1      MFHSIEKYRTPAQVLLGLIALTFVGFGVSTVSHPGADYIVQVGDEKISDHSINNAIQNEQ
                    10        20        30        40        50        60

                    70        80        90       100       110       120
a230-1.pep  ADGGGPSRDAVFQSLLQRAYLKQGAKLMGISVSSEQIKQIIVDDPNFHDANGKFDHALLN
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m230-1      ADGGGPSRDAVFQSLLQRAYLKQGAKLMGISVSSEQIKQIIVDDPNFHDANGKFDHALLN
                    70        80        90       100       110       120

                   130       140       150       160       170       180
a230-1.pep  RYLSQRHMSEDQFVEEIRDQFALQNLVNLVQNGVLVGDAQAEQLIRLTQVNRTIRSHTFN
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m230-1      RYLSQRHMSEDQFVEEIRDQFALQNLVNLVQNGVLVGDAQAEQLIRLTQVNRTIRSHTFN
                   130       140       150       160       170       180

                   190       200       210       220       230       240
a230-1.pep  PDEFIAQVKVSEADLQKFYNANKKDYLLPKAVKLEYVALNLKDFADKQTVSETEVKNAFE
            |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
m230-1      PDEFIAQVKVSEADLQKFYNANKKDYLLPQAVKLEYVALNLKDFADKQTVSETEVKNAFE
                   190       200       210       220       230       240

                   250       260       270       280       290       300
a230-1.pep  ERVARLPANEAKPSFEQEKAAVENELKMKKAVADFNKAKEKLGDDAFNHPSSLAEAAKNS
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m230-1      ERVARLPANEAKPSFEQEKAAVENELKMKKAVADFNKAKEKLGDDAFNHPSSLAEAAKNS
                   250       260       270       280       290       300

                   310       320       330       340       350       360
a230-1.pep  GLKVETQETWLSRQDAQMSGMPENLINAVFSDDVLKKKHNSEVLTINSETAWVVRAKEVR
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m230-1      GLKVETQETWLSRQDAQMSGMPENLINAVFSDDVLKKKHNSEVLTINSETAWVVRAKEVR
                   310       320       330       340       350       360

                   370       380       390       400       410       420
a230-1.pep  EEKTLPFAEAKDAVRQAYIRTEAAKLAENKAKDVLTQLNGGKAVDVKWSEVSVLGAQQAR
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m230-1      EEKTLPFAEAKDAVRQAYIRTEAAKLAENKAKDVLTQLNGGKAVDVKWSEVSVLGAQQAR
                   370       380       390       400       410       420

                   430       440       450       460       470       480
a230-1.pep  QSMPPEAYAELLKAKPANGKPAYVRLIGLPAPVIVEVQAVTPPDDIAAQLPLAKQALAQQ
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m230-1      QSMPPEAYAELLKAKPANGKPAYVRLIGLPAPVIVEVQAVTPPDDIAAQLPLAKQALAQQ
                   430       440       450       460       470       480

                   490       500       510
a230-1.pep  QSANTFDLLIRYFNGKIKQTKGAQSVDNGDGQX
            |||||||||||||||||||||||||||||||||
m230-1      QSANTFDLLIRYFNGKIKQTKGAQSVDNGDGQX
                   490       500       510
```

a231.seq (partial)
```
  1 ATGTCAAAAC GAAAATCCAT AAACCGTCCG TATCAAAAAC CGGCGGAACT
 51 GCCGCCGTTG CAA--TAATC CGCCATTTTA CCGTAAAAAC CGCCGCCTGA
101 ACTTTTTTAT CGNGGCAGAC GGCGGTTGCG CGTCTCCGCA AAAATGCAGG
151 GCGCGCGGTT TTCAGACGGC ATTTGCCGTT CAAAGCCGTG CGGTGTCTTT
201 ACCAAATGCC CAACCATTCG CCCACGGCAT CCATCCAATC CTTATTGCCC
251 CCGCCGCTCC TGCCTGCCCG GCGGTACGCC CACGGCGCTT GCGGATTTTT
301 AGC"TTCCAC AATCCTTTGC GTTCCCTTTC CGCCTGAATT TGAGCGTCGG
351 CATAATCGGC AAAATCCGCC TTATCCTGCT GTTCTTTAGC ATAACTTTTA
401 TAATGCCACG CCGCCCCGTC CTGCACCTGC ATCAGGTTCA AATCGGTTTT
451 GCCGACAGAA ACCTGCGCCA CTTCGCGCTG GTAGCGGTCG GTGTCGAACA
501 CGCGGACGCT GACTTTCCTG CCTTCCGCCG CCGCGCGCAG GTTGTCGCGC
551 GAACGCGTGC CGTAAGCCTG TTTCATCTCC GGCGCGTCGA TATACGCCAT
601 CCGGATTTTG TGTTTCGCGC CGTCGCCGTC GATAACGTGA AGGGTGTCGC
651 CGTCATAGAC TTTGGACACC GTGCCTGTGT AGCGGTGGCC GGATTTCGCC
701 GATGCTCGGC GGCGGGCGGG CGCGTCGGAA CCCGCGTCCC CTGCCGCGCC
751 GAGTACGTCG AGTACGGCAA CCGCCGTCCG CACCGCCTCG CTGCCGTACC
801 CCGTATAACC CAACGCACCC AAAAGCGACA AGGCGACGGG AAGCCATTTC
851 ATGATTTTTT TAATCTGCAT ATTTTTCAAA TGCCGATGCC GTCTGAACAT
901 ATC
```

This corresponds to the amino acid sequence <SEQ ID 844; ORF 217.a>:

a231.pep (partial)
```
  1 MSKRKSINRP YQKPAELPPL QNNPPFYRKN RRLNFFIXAD GGCASPQKCR
 51 ARGFQTAFAV QSRAVSLPNA QPFAHGIHPI LIAPAAPACP AVRPRRLRIF
101 SFPQSFAFPF RLNLSVGIIG KIRLILLFFS ITFIMPRRPV LHLHQVQIGF
151 ADRNLRHFAL VAVGVEHADA DFPAFRRRAQ VVARTRAVSL FHLRRVDIRH
201 PDFVFRAVAV DNVKGVAVID FGHRACVAVA GFRRCSAAGG RVGTRVPCRA
251 EYVEYGNRRP HRLAAVPRIT QRTQKRQGDG KPFHDFFNLH IFQMPMPSEH
301 I
```

m231/a231 98.6% identity in 73 aa overlap
```
                    10        20        30        40        50        60
m231.pep    MSKRKSINRPYQKPAELPPLQNNPPFYRKNRRLNFFIAADGGCASPQKCRARGFQTAFAV
            |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
a231        MSKRKSINRPYQKPAELPPLQNNPPFYRKNRRLNFFIXADGGCASPQKCRARGFQTAFAV
                    10        20        30        40        50        60

                    70
m231.pep    QSRAVSLPNAQPFG
            |||||||||||||:
a231        QSRAVSLPNAQPFAHGIHPILIAPAAPACPAVRPRRLRIFSFPQSFAFPFRLNLSVGIIG
                    70        80        90       100       110       120
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 845>:

g231-1.seq
```
  1 ATGTCAAAAC GAAAATCCAT AAACCGTCCG TATCAAAAAC CGGCGGAACT
 51 GCCGCCGTTG CAAAATAATC CGCCATTTTA CCGTAAAAAC CGCCGCcTGA
101 ACTTTTTTAT CGCGGCAGAC GGCGGTTGCG CGTCTCCGCA AAAATGCAGG
151 GCGCGCGGTT TTCAGACGGC ATTTGCCGTT CAAGGCCGTG CGGTGTCTTT
201 ACCAAATGCC CAACCATTCG CCCACGGAAT CCATCCAATC CTTATTGCCC
251 CCGCCGCTCC TGCCTGCCCG GCGGTACGCC CACGGCGCTT GCGGATTTTT
301 AGCTTTCCAC AATCCTTTGC GTTCCCTTTC CGCCTGAATT TGAGCGTCGG
351 CATAGTCGGC AAAATCCGCC TTATCCTGCT GTTCTTTAGC ATAACTTTTA
401 TAATGCCACG CCGCCCCGTC CTGCACCTGC ATCAGGTTCA AATCGGTTTT
451 GCCGGCGGAT ACCTGCGCCA CTTCGCGCTG ATAGCGGTCG GTTTCAAACa
501 CaCgTaCaat gagtttcgtA ccctccGCCG ccgcgcgCAG GTTGtcgcGC
551 GAACgTGTAC CGTAagcgtg TTtcatctcc GGTGCgtcgA TATACGCCaT
601 cCgAATTTta tGTttcgcgc cgtcgcCgtc gATGACGTGA AGGGtatcGC
651 CgtcATAGAC TTTGGACACC Gtgcctgcgt AGctGTGGCC GGATttcgc
```

This corresponds to the amino acid sequence <SEQ ID 846; ORF 231-1.ng>:

g231-1.pep
```
  1 MSKRKSINRP YQKPAELPPL QNNPPFYRKN RRLNFFIAAD GGCASPQKCR
 51 ARGFQTAFAV QGRAVSLPNA QPFAHGIHPI LIAPAAPACP AVRPRRLRIF
```

```
151  GCGCGCGGTT TTCAGACGGC ATTTGCCGTT CAAAGCCGTG CGGTGTCTTT
201  ACCAAATGCC CAACCATTCG CCCACGGCAT CCATCCAATC CTTATTGCCC
251  CCGCCGCTCC TGCCTGCCCG GCGGTACGCC CACGGCGCTT GCGGATTTTT
301  AGCTTTCCAC AATCCTTTGC GTTCCCTTTC CGCCTGAATT TGAGCGTCGG
351  CATAATCGGC AAAATCCGCC TTATCCTGCT GTTCTTTAGC ATAACTTTTA
401  TAATGCCACG CCGCCCCGTC CTGCACCTGC ATCAGGTTCA AATCGGTTTT
451  GCCGACAGAA ACCTGCGCCA CTTCGCGCTG GTAGCGGTCG GTGTCGAACA
501  CGCGGACGCT GACTTTCCTG CCTTCCGCCG CCGCGCGCAG GTTGTCGCGC
551  GAACGCGTGC CGTAAGCCTG TTTCATCTCC GGCGCGTCGA TATACGCCAT
601  CCGGATTTTG TGTTTCGCGC CGTCGCCGTC GATAACGTGA AGGGTGTCGC
651  CGTCATAGAC TTTGGACACC GTGCCTGTGT AGCGGTGGCC GGATTTCGCC
701  GATGCTCGGC GGCGGGCGGG CGCGTCGGAA CCCGCGTCCC CTGCCGCGCC
751  GAGTACGTCG AGTACGGCAA CCGCCGTCCG CACCGCCTCG CTGCCGTACC
801  CCGTATAACC CAACGCACCC AAAAGCGACA AGGCGACGGG AAGCCATTTC
851  ATGATTTTTT TAATCTGCAT ATTTTTCAAA TGCCGATGCC GTCTGAACAT
901  ATCGGAATCG GATTTCAGAC GGCATCTTAA
```

This corresponds to the amino acid sequence <SEQ ID 850; ORF 231-1.a>:

a231-1.pep

```
  1  MSKRKSINRP YQKPAELPPL QNNPPFYRKN RRLNFFIAAD GGCASPQKCR
 51  ARGFQTAFAV QSRAVSLPNA QPFAHGIHPI LIAPAAPACP AVRPRRLRIF
101  SFPQSFAFPF RLNLSVGIIG KIRLILLFFS ITFIMPRRPV LHLHQVQIGF
151  ADRNLRHFAL VAVGVEHADA DFPAFRRRAQ VVARTRAV[illegible] FHLRRVDIRH
201  PDFVFRAVAV DNVKGVAVID FGHRACVAVA GFRRCSAAGG RVGTRVPCRA
251  EYVEYGNRRP HRLAAVPRIT QRTQKRQGDG KPFHDFFNLH IFQMPMPSEH
301  IGIGFQTAS*
```

a231-1/m231-1   99.0% identity in 309 aa overlap

```
                    10        20        30        40        50        60
a231-1.pep   MSKRKSINRPYQKPAELPPLQNNPPFYRKNRRLNFFIAADGGCASPQKCRARGFQTAFAV
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m231-1       MSKRKSINRPYQKPAELPPLQNNPPFYRKNRRLNFFIAADGGCASPQKCRARGFQTAFAV
                    10        20        30        40        50        60

                    70        80        90       100       110       120
a231-1.pep   QSRAVSLPNAQPFAHGIHPILIAPAAPACPAVRPRRLRIFSFPQSFAFPFRLNLSVGIIG
             |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
m231-1       QSRAVSLPNAQPFAHGIHPILIAPAAPACSAVRPRRLRIFSFPQSFAFPFRLNLSVGIIG
                    70        80        90       100       110       120

                   130       140       150       160       170       180
a231-1.pep   KIRLILLFFSITFIMPRRPVLHLHQVQIGFADRNLRHFALVAVGVEHADADFPAFRRRAQ
             ||||||||||||||||||||||||||||||||||||||||||||:||| |||||||||||
m231-1       KIRLILLFFSITFIMPRRPVLHLHQVQIGFADRNLRHFALVAVGIEHAHADFPAFRRRAQ
                   130       140       150       160       170       180

                   190       200       210       220       230       240
a231-1.pep   VVARTRAVSLFHLRRVDIRHPDFVFRAVAVDNVKGVAVIDFGHRACVAVAGFRRCSAAGG
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m231-1       VVARTRAVSLFHLRRVDIRHPDFVFRAVAVDNVKGVAVIDFGHRACVAVAGFRRCSAAGG
                   190       200       210       220       230       240

                   250       260       270       280       290       300
a231-1.pep   RVGTRVPCRAEYVEYGNRRPHRLAAVPRITQRTQKRQGDGKPFHDFFNLHIFQMPMPSEH
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
m231-1       RVGTRVPCRAEYVEYGNRRPHRLAAVPRITQRTQKRQGDGKPFHDFFNLHIFQMPMPSEH
                   250       260       270       280       290       300

                   310
a231-1.pep   IGIGFQTASX
             ||||||||||
m231-1       IGIGFQTASX
                   310
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 851>:

g232.seq

```
  1  atgatgggca acagcctgat tgaatccggt acgtttgtcg ccatcctgtt
 51  tggtcagatt ttgggaacgg cggttgccgg cgcgccgcct tatattgtcg
101  ggatactggt tttgctggtc gccgtcggag gaacggccgg cagcctgttt
```

```
                10        20        30        40        50        60

                70        80        90        100       110       120
m232.pep   TQIEWNIVRGTKSLLRETVRHKPVFTAIIGISWFWFVGAVYTTQLPTFTQIHLGGNDNVF
           ||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
g232       TQIEWNIVRGTKSLLRETVRHNPVFTAIIGISWFWFVGAVYTTQLPTFTQIHLGGNDNVF
                70        80        90        100       110       120

               130       140       150       160       170       180
m232.pep   NLMLALFSIGIAAGSVLCAKFSXERLMLAWVTVGALGLTVCGLVLVWLTHGHRFEGLNGI
           ||||||||||||||||||||: |||||||||||||| |||||||||||||||||||||||
g232       NLMLALFSIGIAAGSVLCAKFGRERLMLAWVTVGALGSTVCGLVLVWLTHGHRFEGLNGI
               130       140       150       160       170       180

               190       200       210       220       230
m232.pep   FXFLSQGWAYPVMAVMTLIGFFGGFFSVPLYT-VQTAIARFPRP-AVAANNIVNGIFMVS
           | ||||||||||||||||||||||||||||||| :||| ::  |  ||||||||||||||
g232       FWFLSQGWAYPVMA[illegible]TLIGFFGGFFSVPLYTWLQTASSETFRARAVAANNIVNGIFMVS
               190       200       210       220       230       240

          240       250       260       270       280       289
m232.pep   AAVLSAVLLFLFDSISLLYLIVALGNIPLSVFLIKRERRFLGAAAIRKKPX
           |||||||||||||||||||||||||||||:|||||||||||||||||||||
g232       AAVLSAVLLFLFDSISLLYLIVALGNIPLAVFLIKRERRFLGAAAIRKKP
               250       260       270       280       290
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 855>:

```
a232.seq
     1  ATGTACGCTA AAAAAGGCGG TTTGGGACTG GTTAAAAGCC GCCGTTTCGC
    51  ACCTCTTTTC GCTACGCAGT TTCTCGGCGC GTTCAACGAC AATGTGTTCA
   101  AAACCGCGCT GTTTGTGATG ATTGGGTTTT ACGGTTTGGG GCAAAACGGC
   151  TTCCTGCCTG CCGGACAGAT GTTGAACTTG GGCGCGTTGC TGTTTATTTT
   201  GCCGTATTTC CTG[illegible]TTC[illegible]T [illegible]GCTGTCGGG GCAGTTGGGT AACAAAT[illegible]CG
   251  ACAAGGCCGT TTTGGCGCGT TGGGCCAAGG TGCTGGAAAT GATCATTATG
   301  GCGGTGGCGG CATACGGGTT TTATATCCGG TCTGCCCCGC TGCTTTTGGC
   351  GTGTCTGTTT TGCATGGGCG CGCAATCGAC GCTGTTCGGG CCGCTGAAAT
   401  ACGCCATCCT GCCCGATTAT CTCGACGACA AAGAGTTGAT GATGGGCAAC
   451  AGCCTGATTG AATCGGGTAC GTTTGTCGCC ATCCTGTTCG GTCAGATACT
   501  GGGGACTGCG GTGGCAGGTG TACCGCCTTA TATTGTCGGG ATACTGGTTT
   551  TGCTGGTCGC CGTAGGAGGC ACGGTCGGCA GCCTGTTTAT GCCGTCCGTA
   601  CCCGCCAAGG CTGCCGATAC ACAAATTGAG TGGAATATTG TCCGGGGTAC
   651  AAAATCCCTG CTGCGTGAAA CGGTGCGGCA CAAGCCCGTT TTTACCGCCA
   701  TTATCGGTAT TTCGTGGTTT TGGTTTGTCG GCGCGGTTTA TACCACGCAA
   751  CTGCCGACCT TTACCCAAAT CCATCTAGGC GGCAACGACA ATGTTTTCAA
   801  CCTGATGCTT GCCCTGTTTT CCATCGGTAT TGCCGCCGGT TCGGTACTGT
   851  GTGCCAAGTT CAGCAGGGAA CGGCTGAGGT TGGCTTGGGT AACGGTTGGT
   901  GCGTTGGGTT TGACGGTTTG CGGCTTGGTT TTGGTGTGGC TGACGCACGG
   951  ACACCGTTTT GAAGGGCTGA ACGGCATTTT TTGGTTTTTA TCGCAAGGAT
  1001  GGGCATATCC CGTGATGGCG GTGATGACGC TGATCGGCTT TTTCGGCGGA
  1051  TTTTTCTCCG TTCCGCTCTA TACCTGGCTG CAAACCGCCA GTAGCGAGAC
  1101  TTTCCGCGCC CGCGCCGTTG CCGCCAACAA TATCGTTAAC GGTATTTTTA
  1151  TGGTTTCCGC TGCCGTTTTG AGCGCGGTGT TGCTGTTTTT GTTTGACAGC
  1201  ATTTCCTTGT TGTATCTGAT TGTCGCTTTG GGCAATATTC CGTTGTCGGT
  1251  ATTTTTGATT AAGCGCGAAA GGCGGTTTTT AGGCGCGGCG GCAATCAGGA
  1301  AAAAACCTTG A
```

This corresponds to the amino acid sequence <SEQ ID 856; ORF 232.a>:

```
a232.pep
     1  MYAKKGGLGL VKSRRFAPLF ATQFLGAFND NVFKTALFVM IGFYGLGQNG
    51  FLPAGQMLNL GALLFILPYF LFSSLSGQLG NKFDKAVLAR WAKVLEMIIM
   101  AVAAYGFYIR SAPLLLACLF CMGAQSTLFG PLKYAILPDY LDDKELMMGN
   151  SLIESGTFVA ILFGQILGTA VAGVPPYIVG ILVLLVAVGG TVGSLFMPSV
   201  PAKAADTQIE WNIVRGTKSL LRETVRHKPV FTAIIGISWF WFVGAVYTTQ
   251  LPTFTQIHLG GNDNVFNLML ALFSIGIAAG SVLCAKFSRE RLRLAWVTVG
```

```
151  QISATVDRSG LWQAQTPQLF QAGLLHRALA AENLGGITDE ASAVEKLGVR
201  PLLIQGDARN LKLTQPQDAY IVRLLLNAV*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 859>:

```
m233.seq (partial)
  1  ATGAAGCGCA AAAATATCGC GCTGATTCCC GCCGCCGGCA TCGGGGCGCG
 51  TTTCGGTGCG GACAAACCCA AGCAATATGT CGAAATCGGA AGCAAAAACCG
101  TTTTAGAACA TACGATTGGG ATTTTTGAAC GGCATGAGGC CGTCGATTTG
151  ACCGTCGTTG TCGTCTCGCC CGAAGACACG TTTGCCGATA AGGTTCAGAC
201  GGCATTTCCA CAGGTTCGGG TGTGGAAAAA CGGCGGACAG ACCCGCGCCG
251  AAACCGTCCG CAACGGTGTG GCAAAACTGT TGGAAACCGG TTTGGCGGCG
301  GAAACCGACA ATATTCTGGT ACACGATGCC GCGCGTTGCT GCCTGCCGTC
351  TGAAGCTTTG ACGCGGTTGA TAGAACAGGC GGGCAACGCC GCCGAAGGCG
401  GGATTTTGGC AATTCCCATT GCCGATACGC TCAAGTGCGC GGACGGTGGG
451  AACATT....
```

This corresponds to the amino acid sequence <SEQ ID 860; ORF 233>:

```
m233.pep (partial)
  1  MKRKNIALIP AAGIGARFGA DKPKQYVEIG SKTVLEHTIG IFERHEAVDL
 51  TVVVVSPEDT FADKVQTAFP QVRVWKNGGQ TRAETVRNGV AKLLETGLAA
101  ETDNILVHDA ARCCLPSEAL TRLIEQAGNA AEGGILAIPI ADTLKCADGG
151  NI....
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 233 shows 93.4% identity over a 152 aa overlap with a predicted ORF (ORF 233.ng) from *N. gonorrhoeae:*

```
m233/g233

                   10        20        30        40        50        60
m233.pep    MKRKNIALIPAAGIGARFGADKPKQYVEIGSKTVLEHTIGIFERHEAVDLTVVVVSPEDT
            ||||||||||||||:|||||||||||||||||||||||::||||||||||||||||||||
g233        MKRKNIALIPAAGIGVRFGADKPKQYVEIGSKTVLEHVLGIFERHEAVDLTVVVVSPEDT
                   10        20        30        40        50        60

                   70        80        90       100       110       120
m233.pep    FADKVQTAFPQVRVWKNGGQTRAETVRNGVAKLLETGLAAETDNILVHDAARCCLPSEAL
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
g233        FADKVQTAFPQVRVWKNGGQTRAETVRNGVAKLLETGLAAETDNILVHDAARCCLPSEAL
                   70        80        90       100       110       120

                  130       140       150
m233.pep    TRLIEQAGNAAEGGILAIPIADTLKCADGGNI
            :||||||||||||||||:|:|||||  |::|:|
g233        ARLIEQAGNAAEGGILAVPVADTLKRAESGQISATVDRSGLWQAQTPQLFQAGLLHRALA
                  130       140       150       160       170       180
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 861>:

```
a233.seq
  1  ATGAAGCGCA AAAATATCGC GCTGATTCCC GCCGCCGGCA TCGGGGCGCG
 51  TTTCGGTGCG GACAAACCCA AGCAATATGT CGAAATCGGA AGCAAAAACCG
101  TTTTAGAACA TACGATTGGG ATTTTTGAAC GGCATGAGGC CGTCGATTTG
151  ACCGTCGTTG TCGTCTCGCC CGAAGACACG TTTGCCGATA AGGTTCAGAC
201  GGCATTTCCA CAGGTTCGGG TGTGGAAAAA CGGCGGACAG ACCCGCGCCG
251  AAACTGTCCG CAACGGTGTG GCAAAATTGT TGGAAACCGG TTTGGCGGCG
301  GAAACCGACA ATATTCTGGT ACACGATGCC GCGCGTTGCT GCCTGCCGTC
351  TGAAGCTTTG ACGCGGTTGA TAGAACAGGC GGGCAACGCT GCCGAAGGTG
401  GGATTTTGGC AATTCCCGTT GCCGATACGC TCAAGTGCGC GGACGGTGGG
451  AACATTAGTG CAACCGTCGA GCGGACGAGC CTTTGGCAGG CGCAAACGCC
501  GCAGCTTTTC CGCGCCGGGC TGCTGCACCG CGCATTGGCT GCGGAAAACT
551  TGGACGGCAT TACCGATGAA GCGTCCGCCG TGGAAAAATT GGGCATCCGC
601  CCTTTGCTGG TGCAGGGCGA CGCGCGCAAT TTGAAACTGA CGCAGCCGCA
651  GGACGCATAC ATCGTCAGGC TGCTGCTCGA TGCCGTCTGA
```

```
m234.pep   (partial)
       1   ..GAGEYALSNR EIIGFGGTSG YDATLNGKVL DLAIREAVNS LVQAVDNGAW
      51    QPNR*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 234 shows 94.4% identity over a 54 aa overlap with a predicted ORF (ORF 234.ng) from *N. gonorrhoeae:*

```
m234/g234

                                                        10        20        30
m234.pep                                      GAGEYALSNREIIGFGGTSGYDATLNGKVL
                                              ||||||||||||||||||||||||||||||
g234          LGRGKSQIAYAKVALNIVNVNTSEIVYSTQGAGEYALSNREIIGFGGTSGYDATLNGKVL
            140       150       160       170       180       190

                      40        50
m234.pep      DLAIREAVNSLVQAVDNGAWQPNRX
              ||||||||::|||||||||||| |||
g234          DLAIREAVDNLVQAVDNGAWQSNRX
            200       210       220
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 867>:

```
a234.seq   (partial)
       1   AACCGCACCT ATTTGAACGC ATTAAAACAG GAATCCGGCA TTTCCGGCAA
      51   AGCGCATAAC CTGAAAGGCG CAAATTATGT CGNNACCGGC GATGTAACCG
     101   AATTCGGACG CANAGATGTC GGCGATCATC AGCTCTTCGG CATTTTGGGT
     151   CGCGGCAAAT CGCAAATCGC CTATGCAAAA GTGGCTCTGA ATATCGTCAA
     201   CGTCAATACT TCCGAAATCG TCTATTCCGC ACAGGGCGCG GGCGAATACG
     251   CACTTTCCAA CCGTGAAATC ATCGGTTTCG GCGGCACTTC CGGCTACGAT
     301   GCGACTTTGA ACGGCAAAGT TTTAGACTTG GCAATCCGCG AAGCCGTCAA
     351   CAGCCTGGTT CAGGCTGTTG ACAACGGCGC ATGGCAACCC AACCGTTAA
```

This corresponds to the amino acid sequence <SEQ ID 868; ORF 234.a>:

```
a234.pep   (partial)
       1   NRTYLNALKQ ESGISGKAHN LKGANYVXTG DVTEFGRXDV GDHQLFGILG
      51   RGKSQIAYAK VALNIVNVNT SEIVYSAQGA GEYALSNREI IGFGGTSGYD
     101   ATLNGKVLDL AIREAVNSLV QAVDNGAWQP NR*
```

m234/a234 100.0% identity in 54 aa overlap

```
                                                        10                  20
30
m234.pep
GAGEYALSNREIIGFGGTSGYDATLNGKVL

||||||||||||||||||||||||||||||
a234
LGRGKSQIAYAKVALNIVNVNTSEIVYSAQGAGEYALSNREIIGFGGTSGYDATLNGKVL
             50        60        70        80        90       100

                      40        50
m234.pep      DLAIREAVNSLVQAVDNGAWQPNRX
              |||||||||||||||||||||||||
a234          DLAIREAVNSLVQAVDNGAWQPNRX
            110       120       130
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 869>:

```
g235.seq
```

```
m235.pep      YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
              ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
g235          YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVGAVVNQIANSLT
                     130       140       150       160       170       180

                     190       200       210
m235.pep      DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPKX
              ||||||||||||||||||||:|||||||||||||||
g235          DRGYQVSKTAAYNLLSPYSRNGILKGPRFVEEQPKX
                     190       200       210
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 873>:

```
a235.seq
      1  ATGAAACCTT TGATTTTAGG GCTTGCCGCC GTGTTGGCGC TGTCTGCCTG
     51  CCAAGTTCAA AAAGCGCCCG ATTTCGACTA CACGTCATTC AAGGAAAGCA
    101  AACCGGCTTC AATTTTGGTG GTTCCGCCGC TGAACGAATC GCCCGATGTC
    151  AACGGAACAT GGGGTGTACT GGCTTCGACC GCCGCGCCGC TTTCCGAAGC
    201  CGGCTATTAC GTCTTCCCCG CCGCAGTCGT GGAGGAAACC TTCAAACAAA
    251  ACGGCTTGAC CAATGCCGCC GATATTCACG CCGTCCGGCC GGAAAAACTG
    301  CATCAGATTT TCGGCAATGA TGCGGTTTTG TACATTACGG TTACCGAATA
    351  CGGCACTTCA TATCAAATTT TAGACAGCGT GACGACCGTA TCCGCCAAAG
    401  CACGGCTGGT CGATTCCCGC AACGGAAAAG AGTTGTGGTC GGGTTCGGCC
    451  AGCATCCGCG AAGGCAGCAA CAACAGCAAC AGCGGCCTGT TGGGGGCTTT
    501  GGTCAGCGCA GTGGTCAATC AGATTGCCAA CAGCCTGACC GACCGCGGTT
    551  ATCAGGTTTC TAAAACCGCC GCATACAACC TGCTGTCGCC CTATTCTCAC
    601  AACGGCATCT TGAAAGGTCC GAGATTCGTC GAAGAGCAGC CCAAATAA
```

This corresponds to the amino acid sequence <SEQ ID 874; ORF 235.a>:

```
a235.pep
      1  MKPLILGLAA VLALSACQVQ KAPDFDYTSF KESKPASILV VPPLNESPDV
     51  NGTWGVLAST AAPLSEAGYY VFPAAVVEET FKQNGLTNAA DIHAVRPEKL
    101  HQIFGNDAVL YITVTEYGTS YQILDSVTTV SAKARLVDSR NGKELWSGSA
    151  SIREGSNNSN SGLLGALVSA VVNQIANSLT DRGYQVSKTA AYNLLSPYSH
    201  NGILKGPRFV EEQPK*
```

m235/a235   100.0% identity in 215 aa overlap

```
                      10        20        30        40        50        60
m235.pep      MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a235          MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
                      10        20        30        40        50        60

                      70        80        90       100       110       120
m235.pep      AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a235          AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
                      70        80        90       100       110       120

                     130       140       150       160       170       180
m235.pep      YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a235          YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
                     130       140       150       160       170       180

                     190       200       210
m235.pep      DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPKX
              ||||||||||||||||||||||||||||||||||||
a235          DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPKX
                     190       200       210
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 875>:

```
g236.seq
```

```
                                                  10        20        30
m236.pep                                  LHGRTDGFVGAQRLDGGGYRFAGFADCRPF
                                          |:||||:||||||||:||| || |||||||
g236        FRHQQGKAQFFAQSIQIAGHFFRRGNFGFRLQGRTDSFVGAQRLDSGGYCFARFADCRPF
              60        70        80        90       100       110

                    40        50        60        70        80        90
m236.pep    FHQFGFGFFVDGRELVPSMEEDAVXFAAAXDVPRFFAGEAQNRCNQENQTACDVIQGSLC
            |||||||||||||||||||||||| |||| |||||||||||||||||||:| ||:||:|
g236        FHQFGFGFFVDGRELVPSMEEDAVFFAAADDVPRFFAGEAQNRCNQENQAARDVVQGGLR
             120       130       140       150       160       170

                   100       110       120       130       140       150
m236.pep    AAACMAVCFGGVEAVFQDVEVERTQVFRAERNXVFYGKVEXITRIVIACQTLLQLTCQYH
            |||  || |||||||||||||||:|||||||| ||:|:|| |:||| |||||||   ||:
g236        AAAGAAVGFGGVEAVFQDVEVERAQVFRAERNNVFHGEVEGIARIVTACQTLLQPPRQYQ
             180       190       200       210       220       230

                   160       170       180       190       200       210
m236.pep    GVAVDFHHIRLLHGIFNRIKVAQVGKQKAQGIADTAVAFGYAFEDFFGNRQFAAVIGRCR
            |||||||||||||||||||||||:|||:||||||:||||| |||||||||||||||| ||
g236        GVAVDFHHIRLLHGIFNRIKVAQIGKQEAQGIADAAVAFGNAFEDFFGNRQFAAVIGGCR
             240       250       260       270       280       290

                   220       230       240       250      259
m236.pep    PQAQDVCAEFVINLLRCNDVADGFRHFFAFAVDNETMGQQLFIRRATHX
            |||||| |||||::|| :||||||||| |  |::||:||||||| |:|
g236        PQAQDVRAEFVIDFLRRDDVADGFRHFAAVLVNHETVGQQLFIRCASHG
             300       310       320       330       340
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 879>:

```
a236.seq
      1  ATGGCGCGTT TCGCCTTCTC CGCCGACATT CTCTGCACAG CGTTTGCAGA
     51  CGGTTTCATG GCCTGCAACC GCGCCCACAT CGCGGGTGTA GTGCCAGCAG
    101  CGTTCGCATT TTTCACCATC ACTGGCTTTA GCGGCAACGG CAAGTTCGCT
    151  GCCTACTTTC ACTTCTGCTT TAGACACCAG CAAAGCAAAG CGCAATTCTT
    201  CGCCCAAAGC ATTCAGATAG CCGGCCATTT CTTCCGGCGC GGTAATTTCG
    251  GCTTCGGCTT GCAAGGACGA ACCGACGGTT TTGTCGGCGC GCAAAGGCTC
    301  GATGGCGGCG GTTACCGCTT CGCGGGCTTC GCGGATTGCC GTCCATTTTT
    351  TCACCAGTTC GGCTTCGGCT TTTTCGTTGA TGGTCGGGAA CTCGTGCCAA
    401  GTATGGAAAA GCACGCTGTC TTCTGCGCCG CCGCCGATGA TGTCCCACGC
    451  TTCTTCGCCG GTGAAGCACA AAATCGGTGC AATCAAGAGA ACCAGGCTGC
    501  GCGTGATGTG GTACAGGGCG GTTTGCGCGC TGCGGCGGGC GCGGCTGTCG
    551  GCTTTGGTGG TATAGAGGCG GTCTTTCAGG ATATCGAGGT AGAACGCGCC
    601  CAAGTCTTCC GAGCAGAAAG AAACCATTTC TTTCACGGCA AAGTGGAAGG
    651  CATAACGCGG ATAAAAATCA CCGGCAACGC GTTCTTGCAG CCGCCTTGCC
    701  AACACCAAGG CATAGCGGTC GATTTCCACC ATATCCGCCT GTTGCACGGC
    751  ATCTTCAATA GGATTGAAGT CGCTCAAGTT GGCAAACAAA AAGCTCAAGG
    801  TATTGCGGAT ACGGCGGTAG CTTTCGGTTA CGCGCTTGAG GATTTCTTTG
    851  GAAATCGCCA ATTCGCCGCT GTAATCGGTG GATGCCGCCC ACAGGCGCAG
    901  GATGTCCGCG CCGAACTCGT TATACACTTC TTGCGGCGCG ACGACGTTGC
    951  CGATGGATTT CGACATTTTG CGCCCGTTTT GATCCACCAC GAAACCATGG
   1001  GTCAGCAGCT GTTTGTACGG CGCGCGACCC ATTGA
```

This corresponds to the amino acid sequence <SEQ ID 880; ORF 236.a>:

```
a236.pep
      1  MARFAFSADI LCTAFADGFM ACNRAHIAGV VPAAFAFFTI TGFSGNGKFA
     51  AYFHFCFRHQ QSKAQFFAQS IQIAGHFFRR GNFGFGLQGR TDGFVGAQRL
    101  DGGGYRFAGF ADCRPFFHQF GFGFFVDGRE LVPSMEKHAV FCAAADDVPR
    151  FFAGEAQNRC NQENQAARDV VQGGLRAAAG AAVGFGGIEA VFQDIEVERA
    201  QVFRAERNHF FHGKVEGITR IKITGNAFLQ PPCQHQGIAV DFHHIRLLHG
    251  IFNRIEVAQV GKQKAQGIAD TAVAFGYALE DFFGNRQFAA VIGGCRPQAQ
    301  DVRAELVIHF LRRDDVADGF RHFAPVLIHH ETMGQQLFVR RATH*
```

```
101  AFQRPEYRVQ TDISIDGIDH IFTLDAAFGR VNQPTVLIRF DARHRGLLEN
151  LRAGFGSGAG NTQRVIERMK MPGQGIELCA LVHIAVKLLF QLGFIPKLIM
201  TRTVMPLGVF MPLLQLFPML RTDGNRGITA LPIAIDGMFA DAFVHQFDRL
251  QRLLPKPLRL LQADLFFNFP HTAGVVADNL PATPAGRAEA DTCGFQNDGF
301  MSVFRQRQCG TQTAQTAADN AGFGFQTALE FRINSMRINR TKIIRRQIFL
351  KIRANHCVCF IGYIFGRNDT DCRAISSKQK IG*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 883>:

```
m237.seq
   1  ATGCGGGACA AGGTTGGCGG TAATGTCGCA CTCCCCGCCC CACGAATATT
  51  CGATTTTGAC ATCGGCAAGC TGCGGAAAAA CTTTAAGCAT ATCTTGGCGG
 101  ACAAGCTCGG TCATaCGCTC AGGATTGTCG ATAAACTCGT TATCCTTACC
 151  GCCGAAAAGC AGTCTGCCGT CCGCGCTGAG GCGGTAATAA TCCAAAATAT
 201  GGCGGTTGTC GCATACTGCC ATATTGTTAT GGATAAGCCC TTTTGCGCGC
 251  GCCCCCAAGG GTTCGGTCGC AATAATAAAG GTGCTGACAG CAATCGCCTT
 301  GCGTTCCAAA GGCCGGAATA TCGGGTTCAA ACCTGCATAA GTATTGACAG
 351  CATAGACCAC ATTTTTGCAC TCGACGCTGC CTTCGGGCGT GTAAACCAGC
 401  CAACCGTTTT GATGCGGTTC GATGCACGTC ATCGGGGATT GCTCGAAAAT
 451  CTGCGCACCG GCTTCGGCAG CGGCACGAGC GATGCCCAAA GTGTAAGTGA
 501  GCGGATGCAG GTGTCCGGAT AAGGGGTCGA ATTGTGCCCC TTGGTACATA
 551  TCGCTGTCAA GCTGCTGTTT CAACTCGGCT TTATCCCAAA GTTGATAATG
 601  ACTCGCACCG TAATGCCGTT GGGCGTGTTC ATGCCACTGC TGCAACTCTT
 651  CCCAATGCTG CGGACGGACG GCAACCGTGG CATAACCGCG CTGCCAATCA
 701  CAATCGACGG CATGTTTGCG GACGCGTTCG TCCACCAGTT CGACCGCCTG
 751  CAAAGACTGT TGCCAAAACC ATTGCGCCTG CTCCAAGCCG ACCTGTTTTT
 801  CAATTTCCCC CATACCGCAG nCGTAATCGC TGATAACCTG CCCGCCACTC
 851  CGTCCCGACG CGCCGAAACC GATACGCGCG GCTTCCAACA CAACCGTTTC
 901  ATGTCCCTGC TCCGCCAAGG GCAATGCAGT GCACAAACCA CCCAATCCGC
 951  CGCCGATGAT ACAGGTATCG GTTTTCAGAC GGCATTGAAG TTtCGGATAA
1001  ACAGTATGAG GATTAACCGA ACTGAAATAA TAAGAAGGCA GATATTCTTG
1051  AAAATCAGGG CGAATCATTG TGTTTGCTTT ATCAGGTGTA TTTTCGGACG
1101  GAATGATACA GGCTGTCGGG CCATATCGTC CAwACAGAAA ATCGGTTGA
```

This corresponds to the amino acid sequence <SEQ ID 884; ORF 237>:

```
m237.pep
   1  MRDKVGGNVA LPAPRIFDFD IGKLRKNFKH ILADKLGHTL RIVDKLVILT
  51  AEKQSAVRAE AVIIQNMAVV AYCHIVTDKP FCARPQGFGR NNKGADSNRL
 101  AFQRPEYRVQ TCISIDSIDH IFALDAAFGR VNQPTVLMRF DARHRGLLEN
 151  LRTGFGSGTS DAQSVSERMQ VSGXGVELCP LVHIAVKLLF QLGFIPKLIM
 201  TRTVMPLGVF MPLLQLFPML RTDGNRGITA LPITIDGMFA DAFVHQFDRL
 251  QRLLPKPLRL LQADLFFNFP HTAXVIADNL PATPSRRAET DTRGFQHNRF
 301  MSLLRQGQCS AQTTQSAADD TGIGFQTALK FRINSMRINR TEIIRRQIFL
 351  KIRANHCVCF IRCIFGRNDT GCRAISSXQK IG*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 237 shows 86.1% identity over a 382 aa overlap with a predicted ORF (ORF 237.ng) from *N. gonorrhoeae:*

```
m237/g237

                    10        20        30        40        50        60
m237.pep    MRDKVGGNVALPAPRIFDFDIGKLRKNFKHILADKLGHTLRIVDKLVILTAEKQSAVRAE
            ||||||||:||||||||| :||||||||||||||||||| |||||:|||||||| |||||
g237        MRDKVGGNIALPAPRIFDSNIGKLRKNFKHILADKLGHTRRIVDKFVILTAEKQPAVRAE
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m237.pep    AVIIQNMAVVAYCHIVTDKPFCARPQGFGRNNKGADSNRLAFQRPEYRVQTCISIDSIDH
            ||||||||||||||||:||||||| |||| ||||||:|||||||||||||| ||||:|||
g237        AVIIQNMAVVAYCHIVADKPFCARAQGFGGNNKGADGNRLAFQRPEYRVQTDISIDGIDH
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m237.pep    IFALDAAFGRVNQPTVLMRFDARHRGLLENLRTGFGSGTSDAQSVSERMQVSGXGVELCP
```

m237/a237 85.6% identity in 382 aa overlap

```
                        10        20        30        40        50        60
m237.pep        MRDKVGGNVALPAPRIFDFDIGKLRKNFKHILADKLGHTLRIVDKLVILTAEKQSAVRAE
                |||||||||||||||||||||||||||||||||||||||  |||||||||||||||||||
a237            MRDKVGGNVALPAPRIFDFDIGKLRKNFKHILADKLGHTRGIVDKLVILTAEKQSAVRAE
                        10        20        30        40        50        60

                        70        80        90       100       110       120
m237.pep        AVIIQNMAVVAYCHIVTDKPFCARPQGFGRNNKGADSNRLAFQRPEYRVQTCISIDSIDH
                |||||||:||||||||:|||||:| |||  |||||||||||:|| |||:|| ||||:: :
a237            AVIIQNMTVVAYCHIVADKPFCTRAQGFCGNNKGADSNRLALQRLEYRIQTGISIDGVHQ
                        70        80        90       100       110       120

                       130       140       150       160       170       180
m237.pep        IFALDAAFGRVNQPTVLMRFDARHRGLLENLRTGFGSGTSDAQSVSERMQVSGXGVELCP
                |||:||||| |||||||:||:| |  : :||||:|||| :|||:|:|||:: | |:|||
a237            IFAFDAAFGGVNQPTVLIRFNAYHGRMLKNLRTSFGSGAGDAQRVIERMEMPGQGIELCA
                       130       140       150       160       170       180

                       190       200       210       220       230       240
m237.pep        LVHIAVKLLFQLGFIPKLIMTRTVMPLGVFMPLLQLFPMLRTDGNRGITALPITIDGMFA
                |||||||||:|:: ||:|||: ||: |||:|||||:|||||||||||||||||:|:||||
a237            LVHIAVKLLLQFSVIPELIMSCTVIFLGVLMPLLQFFPMLRTDGNRGITALPIAINGMFA
                       190       200       210       220       230       240

                       250       260       270       280       290       300
m237.pep        DAFVHQFDRLQRLLPKPLRLLQADLFFNFPHTAXVIADNLPATPSRRAETDTRGFQHNRF
                ||||||||||||||||||||||:|||||| ||| ||||||||||||||||||||||||||
a237            DAFVHQFDRLQRLLPKPLRLLQTDLFFNFLHTAGVIADNLPATPSRRAETDTRGFQHNRF
                       250       260       270       280       290       300

                       310      .320       330       340       350       360
m237.pep        MSLLRQGQCSAQTTQSAADDTGIGFQTALKFRINSMRINRTEIIRRQIFLKIRANHCVCF
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a237            MSLLRQGQCSAQTTQSAADDTGIGFQTALKFRINSMRINRTEIIRRQIFLKIRANHCVCF
                       310       320       330       340       350       360

                       370       380
m237.pep        IRCIFGRNDTGCRAISSXQKIGX
                |  |||||||||||||| |||||
a237            IGYIFGRNDTGCRAISSKQKIGX
                       370       380
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 887>:

```
g238.seq
      1  atgaatttgc ctattcaaaa attcatgatg ctgttggcag cggcaatatc
     51  gatgctgcat atccccatta gtcatgcgaa cggtttggat gcccgtttgc
    101  gcgatgatat gcaggcaaaa cactacgaac cgggtggcaa ataccatctg
    151  tttggtaatg ctcgcggcag tgttaaaaat cgggtttgcg ccgtccaaac
    201  atttgatgca actgcggtcg gccccatact gcctattaca cacgaacgga
    251  caggatttga aggtgttatc ggctatgaaa cccatttttc aggacacgga
    301  cacgaagtac acagtccgtt cgataatcat gattcaaaaa gcacttctga
    351  tttcagcggc ggcgtagacg gcggttttac cgtttaccaa cttcatcgga
    401  cagggtcgga aatacatccc gcagacggat atgacgggcc tcaaggcggc
    451  ggttatccgg aaccacaagg ggcaagggat atatacagct accatatcaa
    501  aggaacttca accaaaacaa agataaacac tgttccgcaa gcccctttt
    551  cagaccgctg gctaaaagaa aatgccggtg ccgcttccgg ttttctcagc
    601  cgtgcggatg aagcaggaaa actgatatgg gaaaacgacc ccgataaaaa
    651  ttggcgggct aaccgtatgg atgatattcg cggcatcgtc caaggtgcgg
    701  ttaatccttt tttaacgggt tttcaagggg tagggattgg ggcaattaca
    751  gacagtgcgg taagcccggt cacagataca gccgctcagc agactctaca
    801  aggtattaat gatttaggaa atttaagtcc ggaagcacaa cttgccgccg
```

```
151  DYPPPGGARD IYSYYVKGTS TKTKTNIVPQ APFSDRWLKE NAGAASGFFS
201  RADEAGKLIW ESDPNKNWWA NRMDDVRGIV QGAVNPFLMG FQGVGIGAIT
251  DSAVSPVTDT AAQQTLQGIN DLGKLSPEAQ LAAASLLQDS AFAVKDGINS
301  AKQWADAHPN ITATAQTALS AAEAAGTVWR GKKVELNPTK WDWVKNTGYK
351  KPAARHMQTL DGEMAGGNKP IKSLPNSAAE KRKQNFEKFN SNWSSASFDS
401  VHKTLTPNAP GILSPDKVKT RYTSLDGKIT IIKDNENNYF RIHDNSRKQY
451  LDSNGNAVKT GNLQGKQAKD YLQQQTHIRN LDK*
```

Computer analysis of this amino acid sequence gave the following results:

<u>Homology with a predicted ORF from *N. gonorrhoeae*</u>

ORF 238 shows 86.0% identity over a 401 aa overlap with a predicted ORF (ORF 238.ng) from *N. gonorrhoeae:*

```
m238/g238

                       10        20        30        40        50        60
m238.pep      MNLPIQKFMMLFAAAISLLQIPISHANGLDARLRDDMQAKHYEPGGKYHLFGNARGSVKK
              |||||||||||:|||||:|:|||||||||||||||||||||||||||||||||||||||:
g238          MNLPIQKFMMLLAAAISMLHIPISHANGLDARLRDDMQAKHYEPGGKYHLFGNARGSVKN
                       10        20        30        40        50        60

                       70        80        90       100       110       120
m238.pep      RVYAVQTFDATAVSPVLPITHERTGFEGVIGYETHFSGHGHEVHSPFDHHDSKSTSDFSG
              || ||||||||||:|:|||||||||||||||||||||||||||||||||:||||||||||
g238          RVCAVQTFDATAVGPILPITHERTGFEGVIGYETHFSGHGHEVHSPFDNHDSKSTSDFSG
                       70        80        90       100       110       120

                      130       140       150       160       170       180
m238.pep      GVDGGFTVYQLHRTGSEIHPEDGYDGPQGSDYPPPGGARDIYSYYVKGTSTKTKTNIVPQ
              |||||||||||||||||||| ||||||||: || | ||||||||::|||||||| | |||
g238          GVDGGFTVYQLHRTGSEIHPADGYDGPQGGGYPEPQGARDIYSYHIKGTSTKTKINTVPQ
                      130       140       150       160       170       180

                      190       200       210       220       230       240
m238.pep      APFSDRWLKENAGAASGFFSRADEAGKLIWESDPNKNWWANRMDDVRGIVQGAVNPFLMG
              ||||||||||||||||||:||||||||||||:||:||| ||||||:|||||||||||| |
g238          APFSDRWLKENAGAASGFLSRADEAGKLIWENDPDKNWRANRMDDIRGIVQGAVNPFLTG
                      190       200       210       220       230       240

                      250       260       270       280       290       300
m238.pep      FQGVGIGAITDSAVSPVTDTAAQQTLQGINDLGKLSPEAQLAAASLLQDSAFAVKDGINS
              |||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
g238          FQGVGIGAITDSAVSPVTDTAAQQTLQGINDLGNLSPEAQLAAASLLQDSAFAVKDGINS
                      250       260       270       280       290       300

                      310       320       330       340       350       360
m238.pep      AKQWADAHPNITATAQTALSAAEAAGTVWRGKKVELNPTKWDWVKNTGYKKPAARHMQTL
              |:|||||||||||||||||::|||||||||||||||||||||||||||||||||||||||:
g238          ARQWADAHPNITATAQTALAVAEAAGTVWRGKKVELNPTKWDWVKNTGYKKPAARHMQTV
                      310       320       330       340       350       360

                      370       380       390       400       410       420
m238.pep      DGEMAGGNKPIKSLPNSAAEKRKQNFEKFNSNWSSASFDSVHKTLTPNAPGILSPDKVKT
              ||||||||:| ||: :| ::    :: |: :: :  :::::
g238          DGEMAGGNRPPKSI-TSEGKANAATYPKLVNQLNEQNLNNIAAQDPRLSLAIHEGKKNFP
                       370        380       390       400       410

                      430       440       450       460       470       480
m238.pep      RYTSLDGKITIIKDNENNYFRIHDNSRKQYLDSNGNAVKTGNLQGKQAKDYLQQQTHIRN

g238          IGTATYEEADRLGKIWVGEGARQTSGGGWLSRDGTRQYRPPTEKKSQFATTGIQANFETY
             420       430       440       450       460       470
```

```
             ||||||||||||||||||||||||||||||| ||:||||||||||: |||||||||||||
a238         FQGVGIGAITDSAVSPVTDTAAQQTLQGINHLGNLSPEAQLAAATALQDSAFAVKDGINS
                     250       260       270       280       290       300

                     310       320       330       340       350       360
m238.pep     AKQWADAHPNITATAQTALSAAEAAGTVWRGKKVELNPTKWDWVKNTGYKKPAARHMQTL
             |:||||||||||||||||||::|||| ||| ||||||||||||||||||||| ||:! |:||
a238         ARQWADAHPNITATAQTALAVAEAATTVWGGKKVELNPTKWDWVKNTGYKTPAVRTMHTL
                     310       320       330       340       350       360

                     370       380       390       400       410      419
m238.pep     DGEMAGGNKPIKSLP-NSAAEKRKQNFEKFNSNWSSASFDSVHKTLTPNAPGILSPDKVK
             ||||||||:| ||:   || |:    |
a238         DGEMAGGNRPPKSITSNSKADASTQ
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 893>:

```
g239.seq
      1  atgttccacc ataaaggtat tgcccgaaac cggcggatgg aggttttgtt
     51  tttctgccgc cgccctgatc gcttcgtgat tcgccaaacg cgcctgttgc
    101  agcctcattt gcgcataatc ctgctccaag gcgatttcct gtttttttcgc
    151  cttgtccaaa gctgtgaagt tgagcctgta ctggttttgc tgcatcacaa
    201  cggaaaaagc ggaaacgcac accgcaagca gcagaaagaa attcgatttg
    251  ttcattgccg ttcagacgtt tttctctgtt attattccgg tatcggaccg
    301  gcagtccgct ccgccacacg caaaactgcg ctcctcgccc tcgggttggc
    351  ggcaatttcc gcttcacccg gctttaatgc cctgccccacg attttcaggg
    401  gcggatcggg caaatccgct tctctgaccg ccgcccagct cggcaggggc
    451  tcgtgttgcg aatatttttt gacaaactgc ttcacaatgc ggtcttccaa
    501  cgaatggaaa gcaatgaccg ccaaacgccc gccctctttc agacggcaca
    551  tgacctgcgg caataccgcc cctacttctt caagctcgcg gttaataaag
    601  atgcggattg cctggaaggt gcgcgtcgca ggatcctgcc cccgctcgcg
    651  agtacggacg ttttgtgcca cgatctgcgc cagcttgcgg gttgtatcga
    701  ttggactttc cgcccgttgc gcgacaatgg cgcgcacaat ctggcggcta
    751  aaccgctctt caccataa
```

This corresponds to the amino acid sequence <SEQ ID 894; ORF 239.ng>:

```
g239.pep
      1  MFHHKGIARN RRMEVLFFCR RPDRFVIRQT RLLQPHLRII LLQGDFLFFR
     51  LVQSCEVEPV LVLLHHNGKS GNAHRKQQKE IRFVHCRSDV FLCYYSGIGP
    101  AVRSATRKTA LLALGLAAIS ASPGFNALPT IFRGGSGKSA SLTAAQLGRG
    151  SCCEYFLTNC FTMRSSNEWK AMTAKRPPSF RRHMTCGNTA PTSSSSRLIK
    201  MRIAWKVRVA GSCPRSRVRT FCATICASLR VVSIGLSARC ATMARTIWRL
    251  NRSSP*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 895>:

```
m239.seq
      1  ATGCTCCACC ATAAAGGTmy kGCCCGAAAC CGGCkGATGG AGGTTTTGTT
     51  TTTCTGCCGC CGCCCTGATC GCTTCGTGGT TCGCCAAACG CGCCTGTTGC
    101  AGCCTCATTT GCGCATAATC CTGCTCCAAG GCGATTTCCT GTTTTTTCGC
    151  CTTATCCAAA GCTGTGAAAT TGAGCCTGTA CTGGTTTTGC TGCATCACAA
    201  CGGAAAAAGC GGAAACGCAC ACCGCAAGCA GCAGAAGGAA ATTCAATTTG
    251  TTCATTGCCA TTCAGACGTT TTTCTCTGTG ATTGTTCCGG TATCGGACCG
    301  GCAGTCCGCT CCGCCACACG CAAAACCGCA CTTCTCGCCC TCGGATTGGC
    351  GGCAATTTCC GCCTCACCCG GCTTTAATGC CCTGCCCACG ATTTTCAGGG
    401  GCAGCTCGGG CAAATCCGCT TCCCTGaCCG CCGCCCAGCG CGGCAGGGGC
    451  GCGTGTTGCG AATATTTTTT GACAAACTGC TTCACAATGC GATCTTCCAA
    501  CGAATGGAAA GCAATGACCG CCAAACGTCC GCCCTCTTTC AGACGACACA
    551  TGACCTGCGG CAATACTGCC CCTACTTCTT CAAGCTCGCG GTTAATAAAG
    601  ATGCGGACCG CCTGGAAGGT GCGCGTCGCA GGATCCTGCC CCCGCTCGCG
    651  AGTACGGACG TTTTGTGCCA CGATCTGCGC CAGCTTGCGG GTTGTATCGA
    701  TTGGACTTTC CGCCCGTTGC GCAACAATGG CGCGCGCAAT cCGGCGGCTa
    751  AACCGCTCTT cACCATAA
```

This corresponds to the amino acid sequence <SEQ ID 896; ORF 239>:

This corresponds to the amino acid sequence <SEQ ID 898; ORF 239.a>:

```
a239.pep
      1  MLHHKGIARN RRMEVLFFCR RPDRFVVRQT RLLQPHLRII LLQGDFLFFR
     51  LIQSCEVEPV LVLLHHNGKS GNAHRKQQKE IQFVHCHSDV FLCDCSGIGP
    101  AVRSATRKTA LLALGLAAIS ASPGFNALPA IFRGGSGKSA SLTAAQRGRG
    151  ACCEYFLTNC FTMRSSNEWK AMTAKRPPSF RRHMTCGNTA PTSSSSRLIK
    201  MRIAWKVRVA GSCPRSRVRT FCATICASLR VVSIGLSARC ATMARAIWRL
    251  NRSSP*
```

m239/a239 97.3% identity in 255 aa overlap

```
                    10        20        30        40        50        60
m239.pep    MLHHKGXARNRXMEVLFFCRRPDRFVVRQTRLLQPHLRIILLQGDFLFFRLIQSCEIEPV
            |||||| |||| ||||||||||||||||  ||||||||||||||||||||||||||:|||
a239        MLHHKGIARNRRMEVLFFCRRPDRFVVRQTRLLQPHLRIILLQGDFLFFRLIQSCEVEPV
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m239.pep    LVLLHHNGKSGNAHRKQQKEIQFVHCHSDVFLCDCSGIGPAVRSATRKTALLALGLAAIS
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a239        LVLLHHNGKSGNAHRKQQKEIQFVHCHSDVFLCDCSGIGPAVRSATRKTALLALGLAAIS
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m239.pep    ASPGFNALPTIFRGSSGKSASLTAAQRGRGACCEYFLTNCFTMRSSNEWKAMTAKRPPSF
            |||||||||:||||:||||||||||||||||||||·||||||||||||||||||||||||
a239        ASPGFNALPAIFRGGSGKSASLTAAQRGRGACCEYFLTNCFTMRSSNEWKAMTAKRPPSF
                   130       140       150       160       170       180

                   190       200       210       220       230       240
m239.pep    RRHMTCGNTAPTSSSSRLIKMRTAWKVRVAGSCPRSRVRTFCATICASLRVVSIGLSARC
            |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
a239        RRHMTCGNTAPTSSSSRLIKMRIAWKVRVAGSCPRSRVRTFCATICASLRVVSIGLSARC
                   190       200       210       220       230       240

                   250
m239.pep    ATMARAIRRLNRSSPX
            ||||||| ||||||||
a239        ATMARAIWRLNRSSPX
                   250
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 899>:

```
g240.seq
      1  atgatagaag tcatacattt cttcggcgcc gaaacgcgca gacagtttgc
     51  ttgtgccgac gttggacgat ttctgcataa tgccgcgcac atccaaagag
    101  gggtaaacat gggtatcatc gcgcacggga gacggtccga ttttataagg
    151  ctgcgtattc agccgttcgt tcaaatcggt tttgcccgca tccaatgcct
    201  tcgcaatcac gaacggtttg attgccgaac caggttcgat catatcggtt
    251  acggcacggt tgcgccgctg ttcgctgtct gcccggccgg gtctgttggg
    301  atcgtaggcg ggcgtattgg ccaaggcgag gatttccccc gtgcgggcat
    351  ccaaaaccac caccgttccg gcttttgcct gatggtattc gaccgccttg
    401  ttcaactctt cataggccaa ggtctgaatc ctctgatcga gggaaaggat
    451  gatgtctttg ccgttttgcg gtgctttatt gcgcggggag tccaagctgt
    501  ccacaatatt gccctgccgg tcccgcaaaa caacttccgc gccgtcttcg
    551  ccatacaggc tgtcttcaag cgaaagttcc aaaccttcct gacctttgcc
    601  gtcaatatcg gtaaatccga tgacgtgtgc aaacaggttg cccatcgggt
    651  aatggcgttt taa
```

This corresponds to the amino acid sequence <SEQ ID 900; ORF 240.ng>:

```
g240.pep
      1  MIEVIHFFGA ETRRQFACAD VGRFLHNAAH IQRGVNMGII AHGRRSDFIR
     51  LRIQPFVQIG FARIQCLRNH ERFDCRTRFD HIGYGTVAPL FAVCPAGSVG
    101  IVGGRIGQGE DFPRAGIQNH HRSGFCLMVF DRLVQLFIGQ GLNPLIEGKD
    151  DVFAVLRCFI ARGVQAVHNI ALPVPQNNFR AVFAIQAVFK RKFQTFLTFA
```

```
251   GCACGGTTGC GCCGCTGTTC GCTGTCTGCC CGGCCGGGCC TGTTGGGATC
301   GTAGGCGGGC GTATTGGCCA AGGCGAGGAT TTCCCCCGTG CGGGCATCCA
351   AAACCACCAC CGTTCCGGCT TTTGCCTGAT GGTATTCGAC CGCCTTGTTC
401   AACTCTTCAT AGGCCAAGGT CTGAATCCTC TGATCGAGGG AAAGGATGAT
451   GTCTTTGCCG TTTTTCGGGG CTTTATTGCG CGGGGAGTCC AAGCTGTCCA
501   CAATATTGCC CTGCCGGTCC CGCAAAACGA CTTCCGCGCC GTCTTCGCCA
551   TGCAGGCTGT CTTCAAGCGA AAGTTCCAAA CCTTCCTGAC CTTTGCCGTC
601   AATATCGGTA AATCCGATGA CGTGTGCAAA CAGGTTGCCC ATCGGGTAAT
651   GGCGTTTTAA
```

This corresponds to the amino acid sequence <SEQ ID 904; ORF 240.a>:

```
a240.pep
  1   MIEVIHFFGT ETRRQFACAD VGRFLHDAAH IQRGVNMGIA HGRRSDFIRL
 51   RIQPFVQIGF ARIQCLRNHK RFDCRTGFDH IGYGTVAPLF AVCPAGPVGI
101   VGGRIGQGED FPRAGIQNHH RSGFCLMVFD RLVQLFIGQG LNPLIEGKDD
151   VFAVFRGFIA RGVQAVHNIA LPVPQNDFRA VFAMQAVFKR KFQTFLTFAV
201   NIGKSDDVCK QVAHRVMAF*
```

m240/a240 99.1% identity in 219 aa overlap

```
                    10        20        30        40        50        60
m240.pep    MIEVIHFFGTETRRQFACADVGRFLHDAAHIQRGVNMGIAHGRRSDFIRLRIQPFVQIGF
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a240        MIEVIHFFGTETRRQFACADVGRFLHDAAHIQRGVNMGIAHGRRSDFIRLRIQPFVQIGF
                    10        20        30        40        50        60

                    70        80        90       100       110       120
m240.pep    ARIQCLRNHKRFDCRTGFDHIGYGTVAPLFAVCPAGPVGIVGGRIGQGEDFPRAGIQXHH
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
a240        ARIQCLRNHKRFDCRTGFDHIGYGTVAPLFAVCPAGPVGIVGGRIGQGEDFPRAGIQNHH
                    70        80        90       100       110       120

                   130       140       150       160       170       180
m240.pep    RSGFCLMVFDRLVQLFIGQGLNPLIEGKDDVFAVFRGFXARGVQAVHNIALPVPQNDFRA
            |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
a240        RSGFCLMVFDRLVQLFIGQGLNPLIEGKDDVFAVFRGFIARGVQAVHNIALPVPQNDFRA
                   130       140       150       160       170       180

                   190       200       210       220
m240.pep    VFAMQAVFKRKFQTFLTFAVNIGKSDDVCKQVAHRVMAFX
            ||||||||||||||||||||||||||||||||||||||||
a240        VFAMQAVFKRKFQTFLTFAVNIGKSDDVCKQVAHRVMAFX
                   190       200       210       220
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 905>:

```
g241.seq
  1   ATGATAGAAG TCATACATTT CTTCGGCACC GAAACGCGCA GACAGTTTGC
 51   TTGTGCCGAC GTTGGACGAT TTCTGCATGA TGCCGCGCAC ATCCAAAGAG
101   GGGTAAACAT GGGTATCGCG CACGGGAGAC GGTCCGATTT TATAAGGCTG
151   CGTATTCAGC CGTTCGTTCA AATCGGTTTT GCCCGCATCC AATGCCTTCG
201   CAATCACAAA CGGTTTGATT GCCGAACCGG GTTCGATCAT ATCGGTTACG
251   GCACGGTTGC GCCGCTGTTC GCTGTCTGCC CGGCCGGGCC TGTTGGGATC
301   GTAGGCGGGC GTATTGGCCA AGGCGAGGAT TTCCCCCGTG CGGGCATCCA
351   ArACCACCAC CGTTCCGGCT TTTGCCTGAT GGTATTCGAC CGCCTTGTTC
401   AACTCTTCAT AGGCCAAGGT CTGAATCCTC TGATCGAGGG AAAGGATGAT
451   GTCTTTGCCG TTTTTCGGGG CTTTAkTGCG CGGGGAGTCC AAGCTGTCCA
501   CAATATTGCC CTGCCGGTCC CGCAAAACGA CTTCCGCGCC GTCTTCGCCA
551   TGCAAGCTGT CTTCAAGCGA AAGTTCCAAA CCTTCCTGAC CTTTGCCGTC
601   AATATCGGTA AATCCGATGA CGTGTGCAAA CAGGTTGCCC ATCGGGTAAT
651   GGCGTTTTAA
```

This corresponds to the amino acid sequence <SEQ ID 906; ORF 241.ng>:

```
g241.pep
  1   MPTRPTRAAN PPTPTTWLQT AYCPRPPYRP PSVQTHTPHE PASSTCAAKS
```